Journal of Experimental Psychology

ARTHUR W. MELTON, Editor
Department of Psychology, University of Michigan
Ann Arbor, Michigan

DAVID A. GRANT, Associate Editor
University of Wisconsin

WILLIAM K. ESTES, Associate Editor
Indiana University

Consulting Editors

NORMAN H. ANDERSON, *University of California, Los Angeles*
E. JAMES ARCHER, *University of Wisconsin*
FRED ATTNEAVE, III, *University of Oregon*
CLETUS J. BURKE, *Indiana University*
JAMES DEESE, *Johns Hopkins University*
PAUL M. FITTS, *University of Michigan*
FRANK A. GELDARD, *University of Virginia*
JAMES J. GIBSON, *Cornell University*
CLARENCE H. GRAHAM, *Columbia University*
HAROLD W. HAKE, *University of Illinois*
ARTHUR L. IRION, *Tulane University*
HOWARD H. KENDLER, *New York University*
DONALD B. LINDSLEY, *University of California, Los Angeles*
KENNETH MACCORQUODALE, *University of Minnesota*
QUINN MCNEMAR, *Stanford University*
L. STARLING REID, *University of Virginia*
KENNETH W. SPENCE, *State University of Iowa*
DELOS D. WICKENS, *Ohio State University*

CONTENTS

Components of Risk in Decision Making: Probability and Variance Preferences: C. H. COOMBS AND D. G. PRUITT 265

Facilitation of Concept Formation through Mediated Generalization: S. A. MEDNICK AND J. L. FREEDMAN 278

Effect on Visual Threshold of Light Outside the Test Area: I. T. KAPLAN AND H. RIPPS 284

Association among Stimuli and the Learning of Verbal Concept Lists: J. RICHARDSON 290

Figure-Ground Dominance as a Function of Sector Angle, Brightness, Hue, and Orientation: T. OYAMA 299

Effects of Range of Payoffs as a Variable in Risk Taking: J. L. MYERS AND E. SADLER ... 306

Eyelid Conditioned Responses with Various Levels of Anxiety: M. R. BARON AND J. P. CONNOR 310

Local Variations in the Magnitude of a Figural Aftereffect: D. R. MEYER, S. SUKEMUNE, AND R. MYERS 314

Two-Dimensional Tracking with Identical and Different Control Dynamics in Each Coordinate: R. CHERNIKOFF, J. W. DUEY, AND F. V. TAYLOR 318

Instrumental Escape Performance as a Function of the Intensity of Noxious Stimulation: M. A. TRAPOLD AND H. FOWLER 323

Supplementary Report: Delay of Knowledge of Results, Knowledge of Task, and Intertrial Interval: M. R. DENNY, M. ALLARD, E. HALL, AND M. ROKEACH 327

American Psychological Association

Vol. 60, No. 5 November 1960

The JOURNAL OF EXPERIMENTAL PSYCHOLOGY is published monthly, two volumes per year, by the American Psychological Association, Inc., at Prince and Lemon Sts., Lancaster, Pa. The subscription rate per volume is $8.00, or $16.00 annually. Single copies are $1.50. Subscriptions, orders, address changes, and business communications should be addressed to the American Psychological Association, Inc., 1333 Sixteenth St. N.W., Washington 6, D. C.

This Journal publishes original experimental investigations which are intended to contribute toward the development of psychology as an experimental science. Studies with normal human subjects are favored over studies involving abnormal or animal subjects, except when the latter are specifically oriented toward the extension of behavior theory. Experimental psychometric studies and studies in applied experimental psychology or engineering psychology may be accepted if they have broad implications for experimental and theoretical psychology.

Normally, articles exceeding 20 printed pages cannot be accepted. Within this limit, the piecemeal experiment-by-experiment reporting of psychological research is discouraged, and the reporting of the data and conclusions of substantial segments of research efforts is encouraged. Specifically, an integrated series of studies accomplished simultaneously (e.g., most doctoral dissertations) *must* be presented in a single article. Special provision is made for the publication of Supplementary Reports and Replication Reports in articles of not more than 1.5 printed pages (see Editor's Note, this Journal, 1957, 53, 1–2).

Address all articles submitted for publication to the Editor, Arthur W. Melton, Department of Psychology, University of Michigan, Ann Arbor, Michigan. **Manuscripts *must* adhere to the conventions concerning reference citations, preparation of tables and figures, manuscript format, etc., as described in the *Publication Manual of the American Psychological Association.*** When in doubt about practices of the Journal, authors should examine a recent issue.

Articles are published in the order of their receipt, except in rare circumstances. Authors are supplied with 50 free offprints without covers. All of the cost of an author's alterations in galley proof is charged to the author. Priority in publication is given to articles whose authors assume the full cost of publication. In 1960 the publication cost is $30 per printed page.

To save printing costs, supplementary material to articles in this Journal are deposited with the American Documentation Institute. Copies may be obtained by ordering the Document Number given in the footnote reference to ADI. Orders should be sent to ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington 25, D. C. Orders *must* be accompanied by checks or money orders payable to the Chief, Photoduplication Service, Library of Congress.

ARTHUR C. HOFFMAN, Managing Editor
HELEN ORR, Promotion Manager
BEATRICE B. HAMLIN, Editorial Assistant

Second-class postage paid at Lancaster, Pa.

Send address changes to 1333 Sixteenth St. N.W., Washington 6, D. C. Address changes must reach the Subscription Office by the 10th of the month to take effect the following month. Undelivered copies resulting from address changes will not be replaced; subscribers should notify the post office that they will guarantee second-class forwarding postage. Other claims for undelivered copies must be made within four months of publication.

© 1960 by the American Psychological Association, Inc.

Journal of
Experimental Psychology

VOL. 60, NO. 5 NOVEMBER 1960

COMPONENTS OF RISK IN DECISION MAKING: PROBABILITY AND VARIANCE PREFERENCES[1]

C. H. COOMBS AND D. G. PRUITT[2]

University of Michigan

Perhaps the first model for decision making under risk was the theory that an individual maximizes expected value. This theory is dramatically violated by the St. Petersburg paradox (Chernoff & Moses, 1959) which led Bernoulli (1954) to the notion that utility for money is a nonlinear function of money and a theory that individuals maximize expected utility. This theory "explains" too much since such a wide variety of behavior can be accommodated by it that it is difficult to design crucial experiments. If one adds to it a concept of subjective probability and proposes a theory of maximization of subjectively expected utility (Edwards, 1955), the degrees of freedom in the theory are increased, and the measurement difficulties are multiplied. These weaknesses have led the present authors to explore alternative concepts with which to revise the theory of maximization of expected value.

In the current literature there are two concepts which suggest the direction of revision taken in this study. They are known as probability preferences and variance preferences. Quite recently, Edwards (1953) has published experimental work which he interpreted as showing the existence of probability preferences. In brief, he concluded that 50/50 bets were most preferred and bets with .75 probability of winning were avoided. However, his experiment is subject to the criticism that probability differences among his alternatives are confounded with variance differences, so that certain of his results may be equally well interpreted as variance preferences.

The notion of variance preferences is of longer standing. Half a century ago (Fisher, 1906) it was suggested that individuals base their decisions not only

[1] This study was supported in part by a grant from the National Science Foundation, and in part by Project MICHIGAN, a Department of the Army-sponsored project of the University of Michigan in the field of Combat Surveillance administered by the Signal Corps (Contract DA-36-039 ac 78801). We would like to express our appreciation for many helpful comments and criticisms from Patrick Suppes and Jack Sawyer, who are not to be held responsible for the use we made of their suggestions.

[2] Public Health Service Research Fellow of the National Institute of Mental Health.

on expectation but also on the dispersion of the possible outcomes, i.e., for a given expectation an individual may prefer certain amounts of variance over others. That this is a significant variable has recently been urged by Allais (1953), who criticizes the "American School" for disregarding it; more recently, Royden, Suppes, and Walsh (1959) studied the interaction of variance preferences with expected value in the context of measuring the utility of gambling.

Perhaps one of the principal reasons for the long neglect of a concept of variance preferences is that with the use of the expected utility model, any particular variance preference may be equally well interpreted as a utility function on money. In fact, one of the current methods for constructing such utility functions is based upon preferences among bets with the same expectation and equiprobable outcomes but different variance. This method has been most thoroughly developed by Davidson, Suppes, and Siegel (1957) and has also been used by Coombs and Komorita (1958). It becomes difficult to test for the existence of variance preferences after having used them to construct a utility function. It is proposed here, however, that a test can be made which will permit a choice between the variance preference and the expected utility interpretations. The argument is as follows.

Consider a set of n alternatives, each alternative being a gamble with two outcomes, at probabilities of .5, and all gambles having the same objective expected value. If the expected utility model is assumed, then for any preference ordering of these n alternatives, a utility function for money which is monotonically increasing can be found. The number of possible orderings, then, which might be "explained" by a utility function for money is $n!$

From the point of view of the unfolding theory of preferential choice (Coombs, 1952) the amount of variance of each gamble may be considered a point on a continuum of variance and each individual folds this Joint (J) scale at his ideal point corresponding to his preferred amount of variance. The folded J scale is called an I scale and corresponds to his preference ordering on the gambles.

It is convenient to make the following distinction. If the J scale is the same for all individuals at the level of an ordered-metric scale it is called a "quantitative J scale" whereas if it is the same only at the level of an ordinal scale it is called a "qualitative J scale."

If the preference orderings of all *Ss* are generated from a common quantitative J scale, then the variety of orderings which can occur is, at most, one more than the number of midpoints, $\binom{n}{2} + 1$. If they have only the qualitative J scale in common, then the variety of admissible preference orderings is 2^{n-1}. In each case, however, not only is the *number* of admissible preference orderings limited but also the type; they must be only those obtained by folding a J scale between a pair of midpoints—this leads, for example, to the restriction that the least preferred alternative on any I scale may only be one of the two end stimuli on the J scale.

This distinction[3] between a quantitative J scale and a qualitative J scale is relevant here because if the measures on the J scale were actual *objective* variance, the preference orderings would be limited to those generated from a common quantitative J scale. If the preference orderings were generated from a common qualitative J scale one might speak of a scale of subjective variance. For $n \geq 4$, $n! > 2^{n-1} > \binom{n}{2} + 1$. For example, for $n = 6$ there are 720 possible preference orderings, all leading to monotonically increasing utility for money, but only a specified subset of 32 of them may be interpreted as *subjective* variance preferences, and only 16 of these as preferences generated

[3] Another way to formulate the distinction between a quantitative and qualitative J scale is in terms of a cardinal J scale on which an individual's single-peaked preference function descends linearly in both directions from the ideal point (a quantitative J scale) or descends monotonically (a qualitative J scale).

from a scale of the *objective* variances. The variance models are clearly more vulnerable to disproof than the general utility model and it should be possible to distinguish among these alternative explanations experimentally.

Another variable is the possible existence of probability preferences. Any alternative here considered may be completely described by a probability density distribution over the possible outcomes of amounts of money. A distribution may be characterized by a set of parameters such as its moments or functions of them, i.e., each distribution may be mapped into a point in a multidimensional space, and the preferential choices of an individual may induce a transitive ordering on these points.

The moments of a distribution lead to a simple and efficient system for characterizing the stimuli in certain decision problems. The most common and perhaps most useful parameters of a distribution are those obtained from the first three moments, the expected value (E), the variance (V), and the skewness (S). In the special case of alternatives with just two outcomes, a and b, with Probabilities p and q, respectively, these parameters may be calculated as follows:

$$E = pa + qb \qquad [1]$$

$$V = pq(a - b)^2 \qquad [2]$$

$$S = \frac{1 - 2p}{\sqrt{pq}} \qquad [3]$$

From Equation 3 it is evident that skewness is monotone with p and that a probability preference may be equally well interpreted as a preference for skewness. In psychological terms a preference for skewness captures the idea of a preference for "odds," e.g., a distribution extremely skewed to the right represents a "long shot" and a distribution extremely skewed to the left represents an "almost sure thing." An advantage of a concept of preference for skewness includes being able to deal with gambles with more than two outcomes whereas a concept of probability preferences is currently limited to alternatives with just two outcomes.

If two gambles are alike in all respects but expectation it would seem "rational" for the gamble with the higher expectation to be preferred, since more money is preferable to less. If, however, the choice is between two gambles with the same expectation and skewness, differing only in variance, there is no "correct" choice on a priori grounds. The same may be said for gambles which differ only in skewness. Preferences between gambles based on these latter two dimensions reflect individual personality characteristics, which would have significant implications for selection of decision makers and for automatization of the decision process provided these characteristics do not apply solely to the situations in which they were measured.

The concept of variance preference is commonly interpreted psychologically as a utility for risk, i.e., for 50/50 bets with constant expectation, the more variance one prefers the higher the utility for risk. But one might also reasonably maintain that a preference for skewness is also a kind of utility for risk and the question naturally arises: if these exist in the first place, what relation might they have to each other?

The purpose of this study is to test the adequacy of a concept of variance preference as opposed to a nonlinear utility for money, to test for the existence of preference for skewness, and to explore the possibility of an interrelation between variance and skewness preferences if they exist.

METHOD

Subjects.—The *S*s were 99 students at the University of Michigan. There were 43 men ranging in age from 18 to 27 yr. with a median of 20, and 56 women ranging in age from 16 to 47 yr. with a median of 19. All were volunteers, mostly from undergraduate psychology courses.

Stimuli.—In Table 1 are shown the bets used as stimuli. All bets have an expected value of zero. The bets are organized into

TABLE 1

TABLE OF GAMBLES USED AS STIMULI

Variance	1: $S = 1.79$; $p = 1/6$	Set I 2: $S = 1.15$; $p = 1/3$	Set II 3: $S = 0$; $p = 1/2$	Set III 4: $S = -1.15$; $p = 2/3$	5: $S = -1.79$; $p = 5/6$
A: 0.12		1/3 to win \$.50 2/3 to lose \$.25	1/2 to win \$.35 1/2 to lose \$.35	2/3 to win \$.25 1/3 to lose \$.50	
B: 0.30		1/3 to win \$.80 2/3 to lose \$.40	1/2 to win \$.55 1/2 to lose \$.55	2/3 to win \$.40 1/3 to lose \$.80	
C: Set IV 1.00	1/6 to win \$2.25 5/6 to lose \$.45	1/3 to win \$1.40 2/3 to lose \$.70	1/2 to win \$1.00 1/2 to lose \$1.00	2/3 to win \$.70 1/3 to lose \$1.40	5/6 to win \$.45 1/6 to lose \$2.25
D: 5.00		1/3 to win \$3.20 2/3 to lose \$1.60	1/2 to win \$2.25 1/2 to lose \$2.25	2/3 to win \$1.60 1/3 to lose \$3.20	
E: Set V 25	1/6 to win \$11.25 5/6 to lose \$2.25	1/3 to win \$7.10 2/3 to lose \$3.55	1/2 to win \$5.00 1/2 to lose \$5.00	2/3 to win \$3.55 1/3 to lose \$7.10	5/6 to win \$2.25 1/6 to lose \$11.25
F: 100		1/3 to win \$14.20 2/3 to lose \$7.10	1/2 to win \$10.00 1/2 to lose \$10.00	2/3 to win \$7.10 1/3 to lose \$14.20	

five sets. Set I is the column of 6 bets with a constant skewness of $S = 1.15$ (probability of winning of $p = 1/3$). Set II is the column of 6 bets with $S = 0 (p = 1/2)$. Set III is the column of 6 bets with $S = -1.15 (p = 2/3)$. Set IV is the row of 5 bets with a constant variance of \$1.00. Set V is the row of 5 bets with a constant variance of \$25.00. Sets I, II, and III were used to study variance preferences under constant probability treatment; Sets IV and V were used to study skewness (probability) preferences under constant variance treatment. In each of Sets I, II, and III the stimuli are labeled from A to F in order from low to high variance. In each of Sets IV and V the stimuli are labeled from 1 to 5 in order from $p = 1/6$ to $p = 5/6$.

Within each set, the bets were exhaustively paired, producing 65 distinct pairs, as follows: Sets I, II, and III, 15 pairs each; Sets IV and V, 10 pairs each. Each of the 65 pairs was presented 8 times; an additional 56 pairs of bets, not used in this analysis, brought the total to 576 pairs. The stimuli were presented in dittoed booklets which were constructed so that one replication on all pairs was completed before the next was started. Within each replication, the stimuli were presented in a random order.

Instructions.—The *S*s were told to choose, in each pair, the bet they "would prefer to play" and to indicate this decision by making a check in a box beside the bets. They were also told to treat each bet separately, as if it were the only one offered. This was done in the hope of avoiding multiple-bet strategies such as choosing a few of each kind of bet to keep a "balanced portfolio."

To increase interest and make the situation more realistic, two additional instructions were given: (*a*) at the end of the last session some of the items would be picked and played with no exchange of money so that "you can see how you would have come out if we had played for keeps", and (*b*) after that, *S*s would be "given an opportunity" actually to play two of the bets for real money.

Procedure.—The *S*s appeared in a group on three successive evenings. On Day 1 they went through the first two replications; on Day 2, the next three; and on Day 3, the last three. There was considerable difference between *S*s in the amount of time needed to complete the booklets, ranging from 28 to 74 min. on Day 1 and from 9 to 60 min. on Day 2. Nevertheless, all *S*s were paid the standard rate promised them, \$2.00 per session.

At the end of Day 3, 10 pairs were chosen at random and a die thrown for each to determine the outcome of each bet. The *S*s were encouraged to keep track of what they would have won playing the bets they had picked. Then *S*s were given \$.50 credit and asked whether or not they would care actually to play a bet of the form: 1/2 to win \$.25, 1/2 to lose \$.25. This was done twice. (This bet had appeared eight times in the booklets.)

RESULTS

Analysis of Inconsistency of Preference

Purification of the data.—The first step was to eliminate *S*s whose judgments are not significantly different

from random. Miller and Madow (1954) have shown that with n experimentally independent replications on a pair, if $H = 1$, then $2n(1 - H')/\log_2 e$ has a χ^2 distribution with 1 df, where H' is the actual observed uncertainty computed from the proportion of times each member of the pair was chosen in n presentations. In the case of the average observed uncertainty over a number of pairs of alternatives ($\bar{H}$), the statistic $2n(1 - \bar{H})/\log_2 e$ has a χ^2 distribution for which the number of degrees of freedom is equal to the number of pairs of stimuli over which the average is taken.

The level of significance arbitrarily chosen was approximately the 2% level and the numbers of *S*s retained out of the original 99 for each of the five sets of stimuli are I, 92; II, 96; III, 93; IV, 95; V, 94.

The effect of this "purification" of the data was to eliminate 2 *S*s, each of whom was eliminated from three of the stimulus sets, 4 *S*s, each of whom was eliminated from two of them, and 11 *S*s, each of whom was eliminated from one of them. No *S* was dropped from all sets of stimuli.

Intercorrelations and analysis of variance of $\bar{H}$.—The information measure, $\bar{H}$, provides a convenient index number to characterize *S*'s degree of inconsistency over all the pair comparisons in a set of stimuli. Averaging $\bar{H}$ over all *S*s for a set of stimuli provides an index, $\bar{\bar{H}}$, of the average degree of inconsistency associated with a set of stimuli. Table 2 contains the intercorrelations of $\bar{H}$ between the five sets of stimuli, and the bottom row of the table contains the $\bar{H}$ associated with each set. To correct the estimate of $\bar{\bar{H}}$ for bias due to size of sample simply requires adding .09 to each $\bar{\bar{H}}$ in the bottom row of Table 2 (Miller & Madow, 1954). The $\bar{\bar{H}}$ for each set of stimuli is based on all *S*s, prior to "purification." The intercorrelations contained in Table 2, all significant at the .01 level, indicate that individual differences in inconsistency are statistically significant and moderate in magnitude. It is to be noted that the intercorrelations of H between variance and probability preferences tend to be lower than those within either one.

TABLE 2

MEANS AND INTERCORRELATIONS OF $\bar{H}$

	I	II	III	IV	V
II	.55				
III	.67	.43			
IV	.37	.27	.43		
V	.34	.26	.52	.69	
$\bar{\bar{H}}$	.368	.229	.367	.434	.443

Note.—All correlations are significant at the .01 level.

Analyses of variance, based on all *S*s, of the $\bar{H}$ measures for variance preferences under probability treatment and for probability preferences under variance treatment, respectively, were carried out and reveal significantly different $\bar{\bar{H}}$s for variance preferences at different probability levels (more consistency at $p = 1/2$); no significant difference between $\bar{\bar{H}}$s for probability preferences at the different variance levels; and significant individual differences in inconsistency for both variance preferences and probability preferences. A series of t tests revealed that the $\bar{\bar{H}}$ for variance preferences at each probability level are significantly less than the $\bar{H}$ of probability preferences at each variance level, indicating that the inconsistency of variance preferences is less than the inconsistency of probability preferences.

Transitivity of Preferences

If $p(a,b)$ is the probability that Alternative a is preferred to Alternative b then the concept of *weak* stochastic transitivity is defined (Davidson & Marschak, 1959) as follows:

$$p(a,b) \geq .5$$

and

$$p(b,c) \geq .5 \Rightarrow p(a,c) \geq .5$$

To test whether this condition is satisfied, one seeks a permutation of the alternatives such that each element is preferred at least half the time to each alternative to the right of it. If this condition is satisfied by *S*'s pair comparison preferences over a set of stimuli, then a rank order scale from most to least preferred may be constructed, called his I scale. The numbers of *S*s whose preferences satisfied and failed to satisfy this condition for each set of stimuli are given in Table 3.

Clearly weak stochastic transitivity of preferences appears to be the norm. Of the 31 instances in which data failed to satisfy weak stochastic transitivity for some *S* over one of the stimulus sets, 30 of them would satisfy the condition by the reversal of a single 5/3 split in the eight judgments on some one pair of alternatives. So it appears that failure to satisfy weak stochastic transitivity is a consequence of inconsistency and not an inherent intransitivity, which would have been reflected in stochastic intransitivity. This is further borne out by comparing the average *H* of those *S*s who satisfy with those who fail to satisfy weak stochastic transitivity.

TABLE 3

NUMBERS OF *S*S SATISFYING WEAK STOCHASTIC TRANSITIVITY OF PREFERENCES AND THEIR AVERAGE INCONSISTENCY INDEX

Stimulus Sets	Weak Stochastic Transitivity		Stochastic Intransitivity	
	N	*H*	*N*	*H*
Variance Pref.				
I	86	.2952	6	.5085
II	91	.2005	5	.3406
III	85	.3077	8	.5918
Probab. Pref.				
IV	89	.4024	6	.5777
V	88	.4021	6	.6650
Total	439	.3212	31	.5466

Unfolding Analysis of Preferences

Variance preferences.—If a common quantitative J scale of variance exists which generates the preferences of the individuals, then for $n = 6$ there are 15 midpoints and hence 16 segments of the scale, to each of which corresponds a unique preference ordering (Coombs, 1952). Each *S*'s preference orderings were constructed based on weak stochastic transitivity and the preference orderings so obtained for each set of stimuli were unfolded and a best quantitative J scale was constructed—best in the sense that folding it yielded a larger proportion of the data obtained than any other. The results are presented in Table 4 for the variance preferences at the three probability treatments of $p = 1/3, 1/2$, and $2/3$, respectively.

The stimuli are identified by the letters A to F in order from low to high variance. The first column in each table, labeled "I scale number," has the ordinal numbers of the successive segments of the variance scale, numbered sequentially from the low variance end to the high variance end. In the second column are the simply ordered I scales corresponding

TABLE 4

I SCALES FOR VARIANCE PREFERENCES AT $p = 1/3, 1/2$ AND $2/3$

p	I Scale Number	Simply Ordered I Scales		Weakly Ordered I Scales	
		Scales	Cases	Scales	Cases
1/3	1	ABCDEF	30	AB(CD)EF	2
				A(BC)DEF	1
	2	BACDEF	2		
	3	BCADEF	2		
	4	CBADEF	8		
	5	CBDAEF	1		
	6	CDBAEF	2	(CD)BAEF	1
	7	DCBAEF	3		
	8	DCBEAF	1		
	9	DCEBAF	3	(DC)EB(AF)	1
	10	DECBAF	1	(DE)CBAF	1
	11	EDCBAF	5		
	12				
	13				
	14				
	15				
	16	FEDCBA	16		
	Total		74		6
1/2	1	ABCDEF	17		
	2	BACDEF	2		
	3	BCADEF	1	(BC)ADEF	2
	4	CBADEF	5		
	5				
	6	CDBAEF	3	(CD)B(AE)F	1
	7	DCBAEF	4	DCB(AE)F	1
	8	DCBEAF	1	(CD)BEAF	1
	9			D(CE)BAF	2
	10	DECBAF	1		
	11	EDCBAF	7		
	12	EDCBFA	1	EDC(BF)A	1
	13	EDCFBA	1	ED(CF)BA	1
				ED(CF)(BA)	1
	14				
	15			(EF)DCBA	3
	16	FEDCBA	33		
	Total		76		13
2/3	1	ABCDEF	13	A(BC)DEF	1
				(AB)CDEF	3
	2	BACDEF	6	BA(CD)EF	1
	3			BC(AD)EF	1
				(BC)ADEF	1
	4	CBADEF	1	CB(AD)EF	2
	5			C(BD)AEF	1
				CBD(AE)F	1
	6	CDBAEF	1		
	7	DCBAEF	5	DCB(AE)F	2
	8	DCBEAF	1	(CD)BEAF	1
				DCBE(AF)	1
	9			DC(BE)AF	1
	10				
	11				
	12			(DE)CBFA	1
	13				
	14	EDFCBA	1	E(DF)CBA	1
	15	EFDCBA	1	(EF)DCBA	1
	16	FEDCBA	33	FED(BC)A	1
	Total		62		20

to those segments of the J scale, and the next column contains the number of *Ss* from whom that preference ordering was obtained. The interval or I scale number to which an I scale belongs is based on the number of midpoints crossed which is the number of pairwise permutations required to transform the given I scale into the J scale order, ABCDEF. One more than this number is the interval or I scale number. Cells left blank indicate no occurrence of I scales which would correspond to those intervals. Some I scales are only weakly ordered owing to ties between adjacent elements and these are presented in the fourth column, labeled "weakly ordered I scales," in the row of the I scale interval to which they appear likely to correspond, and the tied elements are shown in parentheses. The frequency of their occurrence is in the last column. The I scales that are not simply ordered cannot be assigned to an I scale number with certainty because there are usually two or more intervals to which one may be assigned depending on how the partial order is resolved into a simple order. In each instance *E* made an assignment which seemed a best guess taking the J scale and the stochastic properties of the I scale into account. Alternative assignments rarely differed by more than one interval, so some unreliability has been introduced in order not to sacrifice these data.

In summary, this table shows that at $p = 1/3$ the variance preferences of 80 *Ss* out of 92 can be generated by a common quantitative J scale of variance; at $p = 1/2$, 89/96; at $p = 2/3$, 82/93.

Altogether there are 30 I scales that do not fit this quantitative J scale. Two of these I scales fit the

qualitative J scale but require a different metric. Examining the remaining 28 individually reveals that the deviations are almost entirely due to inconsistency, e.g., 20 of them deviate by only a 5/3 split on one pair.

In brief, the evidence strongly supports the hypothesis that the preference orderings are restricted to the set of orderings generated by a quantitative or qualitative scale of variance, and the great variety of preference orderings admissible under a general utility interpretation does not occur. In addition, the results do not support a theory of decreasing marginal utility, as suggested by Bernoulli (1954), which would predict one and only one preference ordering on the 50/50 bets of Table 4, i.e., the high variance preference FEDCBA, which occurred in only about one-third of the cases.

The predominance of the preference orderings ABCDEF and FEDCBA indicates that most *S*s preferred an extreme variance, either low or high, on any given stimulus set, and about one-third of the cases fall in between. The remarkable absence of alternative I scales in Table 4 with the same I scale number is perhaps due to the increasing magnitude of the successive intervals between stimuli on the variance J scale, which are \$0.18, \$0.70, \$4.00, \$20.00, and \$75.00. The order of the midpoints on the J scale is preserved even in the presence of individual differences in the subjective evaluation of variance. A more nearly equally spaced set of stimuli would have been better for distinguishing between subjective scales of variance.

Probability preferences.—This analysis parallels that of variance preferences except that there are 5 stimuli in a set, hence 10 midpoints which partition the scale into 11 segments to each of which corresponds a unique preference ordering. The preference orderings obtained were unfolded and a best quantitative J scale constructed. The results are presented in Table 5 for probability preferences at low and high variance. The stimuli are identified by the integers 1 to 5 corresponding to probabilities from 1/6 to 5/6, respectively.

The first column has the I scale numbers in order from low probability

TABLE 5

I SCALES FOR PROBABILITY PREFERENCES AT LOW AND HIGH VARIANCE

Variance	I Scale Number	Simply Ordered I Scale		Weakly Ordered I Scales	
		Scales	Cases	Scales	Cases
Low	1	12345	15	1(23)45	2
				123(45)	2
	2				
	3				
	4			3(21)45	1
				321(45)	2
	5	32415	2	32(14)5	1
	6			34(21)5	1
	7			3(42)51	1
	8	34521	2	34(25)1	1
	9				
	10	45321	4	(45)321	2
	11	54321	21		
	Total		44		13
High	1	12345	17	1(23)45	1
				12(34)5	2
				123(45)	1
	2	21345	1	(12)345	1
	3	23145	1		
	4	32145	3	321(45)	1
				32(145)	1
	5			324(15)	1
	6	32451	1		
	7				
	8				
	9				
	10	45321	2	(45)321	2
	11	54321	22	5(43)21	1
				54(32)1	1
	Total		47		12

of winning to high probability of winning. The second column has the simply-ordered scales of probability preferences corresponding to that segment of the J scale, and the next column gives the number of *S*s from whom each preference ordering was obtained. Cells left blank indicate the absence of I scales which would correspond to those segments of the J scale. A number of I scales occurred with ties between adjacent elements. They are presented in the fourth column in the row of the I scale interval to which they appear most likely to correspond. The frequency of their occurrence is in the last column.

In summary, Table 5 shows that at low variance the probability preferences of 57 *S*s out of 95 can be generated by a common quantitative J scale of probability; and at high variance, 59 *S*s out of 94. It is interesting to note that in each instance the ideals are predominantly at the extreme probabilities.

It is also necessary to examine the cases that do not fit a common J scale to see if they are in some sense significant deviations, or whether they are to be attributed to inconsistency. All together there are 73 such cases—61 transitive I scales, 12 intransitive. When these I scales are examined individually it is evident that the predominant reason for their failure to satisfy a folded J scale lies with the stimulus $p = 1/2$, i.e., $S = 0$. It is too much preferred. For example, 8 *S*s had the I scale 13245; 5 *S*s, 31245; 3 *S*s, 35421; 2 *S*s, 53421. In each case Stimulus 3 ($p = 1/2$) is ranked higher than it should be to fold the J scale. If this stimulus is dropped and the I scales based on the remaining four stimuli examined, one finds that at low variance 82 satisfy a folded J scale and 7 do not; at high variance 73 fit a folded J scale and 15 do not. These figures are highly significant and there is little doubt that probability preferences (skewness preferences) exist in the sense of single-peaked preference functions on a J scale.

The overpopularity of symmetric gambles ($p = 1/2$) is perhaps due to the fact that they are regarded as "fair bets," and perhaps also that they are simpler bets, more easily comprehended than some of the others. A naive *S* faced with a choice between two bets, one of which is more complex than the other, might be biased toward the simpler one.

Stability and Interaction of Preference Patterns

As described in the unfolding analysis, to each I scale of preference an I scale number can be assigned corresponding to the ordinal number of the segment of the J scale to which that I scale corresponds. This assignment is unique in the case of a simply-ordered I scale and in the case of ties an I scale number was assigned in the manner previously indicated. These numbers can be regarded as approximate measures of *S*s' ideal points on the J scale and were subjected to correlational and variance analysis.

One question of interest is whether *S*s' ideal points on the J scale shift, relative to each other, under treatment. Evidence of this is contained in the correlations presented in Table 6. All the correlations in the table are significant at the .01 level except the three with the values .17, .18, and .19. It is clear that the probability preferences are quite stable from one to the other level of variance, whereas the variance preferences are much less so under probability treatment. Particularly interesting is the pattern of the correlations between

TABLE 6

INTERCORRELATIONS OF I SCALE NUMBERS OF DIFFERENT J SCALES

Variance and Prob. Preferences	Variance Preferences			Probability Preferences
	$p = \frac{1}{3}$	$p = \frac{1}{2}$	$p = \frac{2}{3}$	Low Variance
$p = 1/2$	.40*			
$p = 2/3$	.18	.67*		
Low Var.	−.45*	.17	.39*	
High Var.	−.41*	.19	.47*	.93*

* Significant at the .01 level.

variance preferences and probability preferences. At low probability, variance preferences are negatively correlated with probability preferences and at high probability they are positively correlated. This will be discussed in conjunction with Table 9 later.

The interrelations of variance and probability preferences may also be studied by analysis of variance. Probability preferences (measured by I scale numbers) are analyzed in a two way ($S \times$ Variance treatment) analysis of variance presented in Table 7. The F tests reveal consistent individual differences in probability preferences over the two variance treatments but no significant effect of the variance treatment on probability preferences. In the latter case, the mean I scale numbers under low and high variance treatment are 6.01 and 6.19, respectively.

In order to study the effect of probability treatment on variance preferences in more detail *Ss* were divided into three groups, A, B, and C, on the basis of their probability preferences. This was feasible because probability preferences were independent of variance treatment and were very stable ($r = .93$). The three groups were formed on the following basis: A was composed of *Ss* with I scale numbers 1 to 4 inclusive ($N = 27$), B was composed of *Ss* with I scale numbers 5 to 9 inclusive ($N = 27$), and C was composed of *Ss* with I scale numbers 10 and 11 ($N = 28$). The few *Ss* (11) whose two I scale numbers under low and high variance treatment did not permit a unique assignment to one of these three groups were omitted. An analysis of variance of the variance preference I scale numbers of these three groups under probability treatment is presented in Table 8. The F tests reveal consistent individual differences in variance preferences over the three probability treatments and a significant effect of the probability treatment on the variance preferences. In the latter case, the

TABLE 7

PROBABILITY PREFERENCES UNDER VARIANCE TREATMENT

	df	*MS*	*F*
Ss	92	30.7482	30.672*
Var.	1	1.6500	1.646
Error	92	1.0025	

* Significant at .01 level.

TABLE 8

VARIANCE PREFERENCES UNDER PROBABILITY TREATMENT BY PROBABILITY PREFERENCE GROUPS

Source	*df*	*MS*	*F*
Between *Ss*	81	65.48	3.096*
Prob. Pref. Groups	2	65.47	1.000
Error	79	65.48	
Between Prob. Treatments	2	285.56	13.502*
Ss × Treatments	162	21.15	
Groups × Treatments	4	308.96	22.852*
Error	158	13.52	

* Significant at the .01 level.

TABLE 9

MEAN VARIANCE PREFERENCES OF PROBABILITY GROUPS UNDER PROBABILITY TREATMENT

Prob. Pref. Groups	Probability Treatment			Means
	1/3	1/2	2/3	
A	9.93	9.41	7.22	8.85
B	6.59	8.52	8.11	7.74
C	3.30	11.95	13.27	9.51
Means	6.57	9.98	9.58	8.71

mean I scale numbers under the three probability treatments of $p = 1/3$, $1/2$, and $2/3$ are, respectively, 6.57, 9.98, and 9.58, signifying that the average ideal variance at $p = 1/2$ and $p = 2/3$ was greater than for $p = 1/3$. When the odds were even or in his favor, the average *S* preferred more variance.

A significant interaction between the probability preference group and probability treatment is observed. The nature of this interaction is made clearer by study of the mean variance preference in each cell of the design (Table 9). This table also reveals what underlies the pattern of correlations between variance preferences and probability preferences presented in Table 6. Probability preference Group A is constituted of *S*s who prefer low probabilities, i.e., long shots, whereas Group C is the other extreme, they prefer the almost sure thing at the small risk of a large loss. Looking at the rows of Table 9, it is evident that the largest mean is in the diagonal, i.e., at the odds or skewness an *S* prefers he prefers more variance. So those *S*s who prefer low probabilities tend to choose high variance bets at those odds. This is the reason the pattern of correlations between probability preferences and variance preferences in Table 6 ranged from significantly negative to significantly positive. It is interesting to note that for fair bets (the column with $p = 1/2$) those who prefer more skewness (Groups A and C) prefer higher variance bets than do *S*s in Group B who prefer fair bets.

DISCUSSION

The results of this experiment support the hypothesis that variance preferences[4] exist and are generated by folding a joint scale. However, because of the flexibility of the expected utility model, it is possible to find, for each variance preference ordering that folds the joint scale, a nonlinear utility function of money which explains the ordering equally well. The fact that the many possible preference orderings admissible under the utility interpretation do not occur, but that instead only those admissible under a variance preference interpretation occur, is not conclusive. The argument of those who would still prefer a utility interpretation might be as follows: the utility functions corresponding to those preference orderings which did not occur may be, in some sense, unreasonable, in spite of the fact that they are still monotonically increasing with money.

The problem which this argument raises is a mathematical one—characterizing the classes of utility functions which would correspond to various preference orderings on variance. It may be, for example, that those utility functions corresponding to preference orderings on a variance scale have fewer inflection points than those utility functions corresponding to preference orderings outside of the set of variance preferences. If one agrees that utility functions for money with many inflection points are unreasonable, and that the reasonable ones are just those which

[4] The term "variance preference" is used in the generic sense of preference for dispersion. As far as these data are concerned, any measure of dispersion which is monotonic with variance is equally admissible as an hypothesis.

correspond to preference orderings which can otherwise be interpreted as variance preferences, then from the data in this experiment, there is less ground for a choice between these two interpretations. If, however, there exists a subset of preference orderings, admissible under a utility interpretation and not admissible under a variance preference, but corresponding to utility functions with at most one inflection point, then this argument for a utility interpretation is defeated. The mathematical problem is still unsolved.

The fact that the variance preferences of *Ss* shift under probability treatment in a way related to their skewness preferences must also put constraints on the admissible utility functions and this might provide a further basis for choosing between these interpretations.

The concept of probability preferences, reformulated as a concept of skewness preferences, is also given support by this study. It appears that individual skewness preferences exist which are very stable over shifts in the level of variance, from a variance of \$1.00 to a variance of \$25.00.

It appears that a psychological concept of utility for risk is not to be interpreted simply as a preference for variance or a preference for skewness but rather as some not very simple function of both of these. There is clear evidence of interaction between these two variables. Speculation from the results would suggest that skewness plays a dominant role, and as a bet deviates from *S*'s preferred odds, he prefers less variance.

Before a model can be constructed in this area, the relation of variance and skewness preferences to expected value (the first moment) must also be studied. An experiment by Royden, Suppes, and Walsh (1959) is closely related to this problem and their findings suggest that variance preference patterns change in a rather complex fashion as expected value varies. The higher moments of the distributions that describe gambles, such as kurtosis, also remain to be investigated.

Summary

An alternative is suggested to the maximization of expected utility model: A gamble can be characterized by the expectation, variance, and skewness of its probability density distribution over money; and while *S* will always prefer more expectation (other things being equal), he may have a utility for risk which will exhibit itself as a preference for certain amounts of variance and skewness. While a variance preference is isomorphically interpretable as a utility function for money, there are many fewer admissible preference orderings under the former interpretation than the latter and this provides a basis for experimental test.

A study was carried out on bets which were constant in expectation but which varied in skewness or variance with the other variable held constant. Replicated pair comparisons permitted measurement of inconsistency of choice behavior, testing stochastic transitivity, construction of the stochastically dominant preference ordering, and an unfolding analysis of these preference orderings. The principal conclusions were as follows:

1. Degree of inconsistency of preferences is a significant characteristic of *Ss*.
2. Weak stochastic transitivity of preferences is satisfied by most *Ss*.
3. Variance preference orderings were confined almost exclusively to the admissible set generated by folding a variance scale. Instances of the greater variety admissible under a utility interpretation were rare. Approximately one-third of the *Ss* preferred low variance, one-third high variance, and one-third preferred intermediate degrees of variance.
4. Except for the inflated popularity of fair bets ($S = 0$, i.e., $p = 1/2$), particularly for women, the existence of skewness preferences in the sense of folding a probability scale was suggested with the great majority of *Ss* having their ideals at one end or the other of the skewness scale.

References

Allais, M. Le comportement de l'homme rationnel devant le risque. Critique des postulate et axioms de l'école americaine. *Econometrica*, 1953, **21**, 503–546.

Bernoulli, D. Exposition of a new theory on the measurement of risk. *Econometrica*, 1954, **22**, 23–36.

CHERNOFF, H., & MOSES, L. E. *Elementary decision theory.* New York: Wiley, 1959.

COOMBS, C. H. A theory of psychological scaling. Ann Arbor: Univer. Michigan Press, 1952. Eng. Res. Inst. Bull. No. 34.

COOMBS, C. H., & KOMORITA, S. Measuring utility of money through decision. *Amer. J. Psychol.*, 1958, **71**, 383–389.

DAVIDSON, D., & MARSCHAK, J. Experimental tests of a stochastic decision theory. In C. W. Churchman & P. Ratoosh (Eds.), *Measurement: Definitions and theories.* New York: Wiley, 1959. Ch. 13.

DAVIDSON, D., SUPPES, P., & SIEGEL, S. *Decision making.* Stanford: Stanford Univer. Press, 1957.

EDWARDS, W. Probability preferences in gambling. *Amer. J. Psychol.*, 1953, **66**, 349–364.

EDWARDS, W. The prediction of decisions among bets. *J. exp. Psychol.*, 1955, **50**, 201–214.

FISHER, I. *The nature of capital and income.* New York: Macmillan, 1906.

MILLER, G. A., & MADOW, W. G. On the maximum likelihood estimate of the Shannon-Wiener measure of information. *USAF Cambridge Res. Cent. tech. Rep.*, 1954, No. 54–73.

ROYDEN, H. L., SUPPES, P., & WALSH, K. A model for the experimental measurement of the utility of gambling. *Behav. Sci.*, 1959, **4**, 11–18.

(Received October 16, 1959)

Journal of Experimental Psychology
1960, Vol. 60, No. 5, 278–283

FACILITATION OF CONCEPT FORMATION THROUGH MEDIATED GENERALIZATION [1]

SARNOFF A. MEDNICK [2] AND JONATHAN L. FREEDMAN [3]

Harvard University

Underwood (1952) has suggested a method for the study of concept formation which assumes that the attainment of a concept calls for the perception of a relationship between concept instances. The perception of this relationship is in large part dependent on the probability of the occurrence of the relevant associative responses to the concept instances. This probability is termed response dominance. The mean response dominance of all instances representing a particular concept is termed dominance level.

Underwood and Richardson (1956b) have shown that the ease of attainment of a concept is directly related to dominance level. One implication of these results and of the theory in general is that it should be possible to facilitate the attainment of a particular concept by increasing the strength of the associations between the concept instances and the correct concept response. Norms giving the response dominance of conceptual responses to 213 instances have been reported by Underwood and Richardson (1956a). However, it seems likely that by certain experimental manipulations these response probabilities could be temporarily altered. If this alteration took the form of strengthening certain chosen associations, it should result in the facilitation of the attainment of the chosen concept.

The present experiment employs a technique involving mediated generalization to increase the strength of response dominance and to facilitate concept formation. This technique is based in part on the work of Russell and Storms (1955) who attempted to demonstrate that mediated generalization would facilitate paired-associate learning. They presented *S*s with two paired-associate lists, the first of which was designed to facilitate the learning of the second. For example, one pair in the first list was "cef-stem"; a pair in the second list was "cef-leaf." "Stem" elicits the associate, "flower," which in turn elicits "leaf." The assumption was that when "cef-stem" is learned, words associated with "stem" occur implicitly and the associations "cef-flower" and "cef-leaf" are also being implicitly learned. Thus the pair "cef-leaf" should be easier to learn than if *S* had not had experience with "cef-stem."

In the present experiment *S*s are also presented with two lists, the first of which is designed to facilitate the mastering of the second. However, instead of two paired-associate lists, one paired-associate list and one concept-formation list are used. By constructing the lists in accordance with the mediated generalization paradigm we attempted to facilitate the attainment of selected concepts.

METHOD

Subjects.—The *S*s were 30 undergraduate volunteers from Harvard and Radcliffe Colleges. There were 7 men and 23 women.

[1] This study was partially supported by a grant from the National Science Foundation (NSF-G 3855).

[2] Now at the University of Michigan.

[3] Now at Yale University.

Design.—The design of the experiment is best explained by reference to Table 1. There are two paired-associate lists (I and II) and a concept-formation list. Paired-Associate List I is designed to facilitate the attainment of the concept "white" in the concept-formation list. Prior learning of Paired-Associate List II is designed to facilitate the attainment of the concept "soft" in the concept formation list.

The method of facilitation might best be explained by example. Thus in order to facilitate the production of the concept response "white" to the concept instance "sugar" (Table 1) prior pairing of the words "sugar-black" is given in Paired-Associate List I. It is assumed that the word "black" elicits the implicit response "white" causing the word "white" to occur in contiguity with the word "sugar," strengthening the association "sugar-white." When *S* is subsequently presented with the instance "sugar" in the concept-formation task the concept "white" will be attained more easily than it would have been if no pretask training had been given. That is, mediated generalization strengthens the association" sugar-white"; strengthening this association raises the response dominance of the concept instance "sugar" (since response dominance is simply the strength of association between "sugar" and "white"); raising the response dominance should facilitate the production of the concept response "white" to the concept instance "sugar." Since all of the instances of the concept "white" are facilitated in this manner by relevant associate pairs in List I we would predict that prior experience with List I would facilitate attainment of the concept "white."

For 15 *S*s (3 men and 12 women) "white" was facilitated and "soft" was not facilitated (List I). For the other 15 *S*s "soft" was facilitated and "white" was not facilitated (List II). Thus we could sum across the facilitated concepts ("white" preceded by learning of List I and "soft" preceded by learning of List II) and sum across the nonfacilitated concepts ("soft" preceded by learning of List I and "white" preceded by learning of List II) and compare performance on facilitated and nonfacilitated concepts knowing that any resultant differences could not be due to unique properties of the specific concepts or concept instances employed. In order to further insure that the facilitated and nonfacilitated concepts differed only with respect to the independent variable, the *responses* were kept the same in the non-

TABLE 1

PAIRED-ASSOCIATE AND CONCEPT-FORMATION LISTS

Paired-Associate List I

Pair	Relevant Assoc.
sugar-black	black-white (.751)[a]
hailstone-sheep	sheep-white (.009)
button-bread	bread-white (.015)
cabbage-salt	salt-white (.006)
pup-afraid	none
grape-swift	none
peach-beautiful	none
cradle-anger	none
anchor-king	none
camel-dream	none
forest-command	none
barrel-lion	none

Concept-Formation List

Instance	Response Dominance	Concept
sugar	11	"white"
hailstone	9	(M = 7.50)[b]
button	5	
cabbage	5	
pup	12	
grape	7	"soft"
peach	5	(M = 8.75)[b]
cradle	11	
anchor	5	
camel	14	"large"
forest	12	(M = 9.25)[b]
barrel	6	

Paired-Associate List II

Pair	Relevant Assoc.
sugar-afraid	none
hailstone-anger	none
button-beautiful	none
cabbage-swift	none
pup-hard	hard-soft (.674)[a]
loud-soft	loud-soft (.541)
peach-smooth	smooth-soft (.206)
cradle-comfort	comfort-soft (.069)
anchor-king	none
camel-dream	none
forest-command	none
barrel-lion	none

[a] Probability of occurrence of response term (e.g., "white") to stimulus term (e.g., "black").
[b] Mean response dominance of concept response (e.g., "white").

facilitated pairs of the two paired-associate lists.

The third concept, "large," was included as a buffer concept in order to increase the difficulty of the concept list. Since its appearance was not counterbalanced with respect to facilitation it could not be used as a control for facilitation.

Materials.—The details of the construction of the lists follow. In making the paired-associate lists 16 words were first selected from the Minnesota Kent-Rosanoff Word Association Test norms (Russell & Jenkins, 1954). Of these 16 words, four have "white" as a common response, four have "soft," and eight have neither "white" nor "soft," as a common response. The probability of each of these associations is given in Table 1. In Paired-Associate List I the four words to which "white" is a common response were paired (as responses) with the four words that are concept instances of "white" in the concept-formation list, while the words that are concept instances of "soft" were paired with neutral words. Paired-Associate List II reversed the procedure: words that are concept instances of "soft" in the concept-formation list were paired with words which elicit "soft" as an associate, while concept instances of "white" were paired with neutral words. All instances of the concept "large" were paired with neutral words in both lists.

As is indicated in Table 1 the words designed to facilitate the attainment of the concept "soft" have a higher mean probability of response than those designed to facilitate the attainment of the concept "white." Due to the limited number of appropriate words available in the Minnesota norms, this inequality could not be avoided. However, the results of a study by Freedman and Mednick (1958) indicate that the polarization effect of the extremely strong association between "black" and "white" might be expected to offset this difference to some extent. In any event, since the experiment is concerned chiefly with a comparison of the facilitated and nonfacilitated concepts, and the concepts "soft" and "white" appear equally in both conditions, any difference in difficulty level which might exist would not affect the major results.

The concept-formation list (Table 1) contains three concepts ("white," "soft," and "large") each of which is represented by four nouns taken from the list of nouns for which Underwood and Richardson (1956a) have ascertained the dominance level of various responses. Based on these norms the means of response dominance of these concepts were made approximately equal. The dominance level was purposely kept quite low because it was thought that any facilitating effect would be more apparent if the lists were difficult. Also, this difficulty level of the concept list would produce a wider range of scores.

Various factors restricted the selection of concepts for the concept list: (*a*) it was necessary to choose concepts that could be found in the Minnesota Word Association norms as responses to the standard word association test; (*b*) it was important to minimize the amount of interference among the various concept instances (e.g., instances such as "rabbit" to which both "white" and "soft" were possible responses had to be avoided); (*c*) the concepts used had to be represented by four very low-dominant instances that could be approximately equalized in terms of mean and variance of dominance. The concepts "white" and "soft" were the only ones found that met all of these requirements.

Procedure.—Each *S* was presented with one of the paired-associate lists and was required to learn it to a criterion of two successive perfect trials. After a slight pause during which tapes were changed, *S* was presented with the concept-formation instructions. It was explained to *S* that the list contained 12 words that could be placed in 3 groups of 4 words each, and that all of the 4 words in each of these groups could be described by the same adjective. The *S* was required to respond to each word with an adjective that he thought described that word. It was made clear that the adjective must describe a physical quality of the noun. When *S* responded *E* said "right" or "wrong." The *S*'s task was to discover the three correct adjectives and to apply them to the correct nouns. The test was continued to a criterion of one perfect trial or was terminated at 18 trials. The procedure for the concept formation task has been described in more detail elsewhere (Underwood & Richardson, 1956a).

Both the paired-associate lists and the concept-formation list were presented at a 4-sec. rate in three different orders on a Gerbrands memory drum. There was an interval of 8 sec. between trials except for an interval of 20 sec. (due to a blank interval on the tape) after every three trials of the concept-formation list. To control for any effect of order of presentation of the particular words the three orders of the concept-formation list and the paired-associate lists were

the same for all *Ss*. The orders of presentation of the *stimulus* words were kept the same in the two paired-associate lists.

Results and Discussion

The data which were subjected to analysis were the number of trials to one perfect trial on a concept (this meant giving the correct concept response to all four instances of a concept on the same trial) and the number of errors made on each concept in the entire course of the experiment. As noted above, the concept "large" was omitted from this analysis. All *Ss* failing to attain a concept in the 18 allotted trials were given a score of 19. Facilitated concepts ("white" preceded by learning List I, "soft" preceded by learning List II) were compared with nonfacilitated concepts ("soft" preceded by learning List I, "white" preceded by learning List II).

Table 2 presents the data for the number of *Ss* who attained facilitated and nonfacilitated concepts on successive blocks of trials. It can be seen that the facilitated concepts were solved earlier and that fewer *Ss* failed to solve them. No *S* failed to solve the facilitated concept whereas seven *Ss* failed to solve the nonfacilitated concepts. The mean number of trials taken to solve each concept was 7.17 for the facilitated concepts and 12.1 for the nonfacilitated concepts, a significant difference ($t = 3.67$, 29 *df*, $P < .001$). The mean numbers of errors on the facilitated and nonfacilitated concepts were 16.13 and 36.77, respectively. This difference was significant ($t = 3.4$ 29 *df*, $P < .005$).

TABLE 2

Number of *Ss* Attaining Concepts on Successive Blocks of Trials

Concept	Trial on Which Concept Attainment Took Place				Failed to Attain	Total
	1–4	5–8	9–12	13–18		
Facilitated	7	16	4	3	0	30
Nonfacilitated	0	10	7	6	7	30
Buffer	3	13	5	2	7	30

Although it is evident that facilitation occurred, there is still the question of how the paired-associate learning facilitated concept attainment. The results of a study by Freedman and Mednick (1958) suggest that the facilitation could have operated by making the first correct response to a concept occur earlier than it would have without the paired-associate learning. If the first correct response to facilitated concepts is given before the first correct response to nonfacilitated concepts, the former would be easier to attain than the latter. The mean trial on which the first correct response was given was 3.5 for facilitated concepts and 8.4 for nonfacilitated concepts. This difference was found to be significant ($t = 2.87$, 29 *df*, $P < .005$).

It is also possible that the facilitated concepts were easier to attain once the correct response was given. That is, the fact that the first correct response to facilitated concepts was given earlier might not be the only explanation of the facilitation; the training with the paired-associate lists might also have enabled *Ss* to apply the correct response (once it was given to one instance of a concept) more easily to the facilitated concepts than to the nonfacilitated concepts. The results show that the facilitated concepts were attained an average of 4.3 trials after the first correct response was given; the figure for nonfacilitated concepts is 3.7 trials. Thus, once the correct response was discovered the nonfacilitated concepts were actually attained somewhat more quickly than the facilitated concepts (but not significantly so). These results (i.e., the facilitated concepts being attained no more quickly once the correct response

was known) can be explained by the fact that usually by the time the nonfacilitated concept was discovered the facilitated concept had already been attained. In addition, as might be predicted from the dominance levels, the buffer concept "large" was on the average attained earlier than the nonfacilitated concepts (mean number of trials to attainment of large was 10.2). This means that in many cases when *S* made his first correct response to the nonfacilitated concept he had already attained the other two concepts (facilitated and buffer). At this point only four instances remained to which the correct response was not known. Thus there was relatively little difficulty in applying the correct response to the remaining instances.

A possible objection to the present experiment might be that *S*s gave the correct response to the facilitated concepts more quickly because the paired-associate learning made them aware of this response. If *S* were conscious of the relationship between the paired-associate list and the concept-formation task, he might be able to discover the correct response to the facilitated concept by looking for the relationship between particular words on both lists (e.g., *S* could question the significance of the pair "sugar-black," realize that "black" is associated with "white," and thus discover the correct response to "sugar"). In other studies demonstrating mediated association or mediated generalization (Bugelski & Scharlock, 1952; Russell & Storms, 1955), it was found that *S*s were unaware of the process of mediation. Immediately following the procedures of the present experiment all *S*s, except for the first four *S*s, were questioned intensively about any observed relationship between the paired-associate and concept-formation lists. Of the 26 *S*s interviewed, none expressed awareness of the relationship between the lists. It seems likely that the facilitation occurred "unconsciously" by strengthening of the appropriate mediating associations.

A variable that may prove to be significant is the strength of the association between the facilitating words (e.g., "black-white" is a much stronger association than "sheep-white"). Although the present study was not designed to investigate this variable, there was some indication in the data that stronger associations have a greater facilitating effect.

It might be asked whether or not it is possible to interfere with the attainment of concepts in a concept-formation test. One possible method for accomplishing this is to strengthen competing responses. For example, common responses to "grape" are "round," "juicy," "purple," and "soft." If "grape" is an instance of the concept "soft," strengthening the associations between "grape" and "juicy," "purple," or "round" should interfere with the attainment of the concept "soft" by making competing responses occur and thus cause the correct response (i.e., "soft") to be retarded. Underwood (1957) has shown that concepts represented by instances with strong competing responses are in fact more difficult to attain than concepts represented by instances that do not have strong competing responses. Therefore, by employing the paired-associate, mediated generalization method used in the present study, to strengthen interfering (i.e., competing) responses, it should be possible to make concept formation more difficult.

Summary

Thirty *S*s learned one of two paired-associate lists of 12 pairs and were then presented with a list of 12 nouns and were instructed to discover into what three groups the nouns could be divided and what adjectives could describe each group. The paired-associate lists were each designed by a method based on mediated generalization to facilitate a concept in the concept-formation list. The facilitated concepts were attained more quickly and with fewer errors than the nonfacilitated concepts. An explanation of the process of facilitation is offered and is supported by an analysis of the results.

References

Bugelski, B. R., & Scharlock, D. P. An experimental demonstration of unconscious mediated generalization. *J. exp. Psychol.*, 1952, **44**, 334–338.

Freedman, J. L., & Mednick, S. A. Ease of attainment of concepts as a function of response dominance variance. *J. exp. Psychol.*, 1958, **55**, 463–466.

Russell, W. A., & Jenkins, J. J. The complete Minnesota norms for responses to 100 words from the Kent-Rosanoff Word Association Test. *ONR tech. Rep.*, 1954, No. 11.

Russell, W. A., & Storms, L. H. Implicit verbal chaining in paired-associate learning. *J. exp. Psychol.*, 1955, **49**, 287–293.

Underwood, B. J. An orientation for research on thinking. *Psychol. Rev.*, 1952, **59**, 209–220.

Underwood, B. J. Studies in distributed practice: XV. Verbal concept learning as a function of intralist interference. *J. exp. Psychol.*, 1957, **54**, 33–40.

Underwood, B. J., & Richardson, J. Some verbal materials for the study of concept formation. *Psychol. Bull.*, 1956, **53**, 84–95. (a)

Underwood, B. J., & Richardson, J. Verbal concept learning as a function of instructions and dominance level. *J. exp. Psychol.*, 1956, **51**, 229–238. (b)

(Received October 6, 1959)

Journal of Experimental Psychology
1960, Vol. 60, No. 5, 284–289

EFFECT ON VISUAL THRESHOLD OF LIGHT OUTSIDE THE TEST AREA [1]

IRA T. KAPLAN [2] AND HARRIS RIPPS

Postgraduate Medical School, New York University

The luminance threshold of a test field can be lowered or raised by the presence of an inducing field. When the inducing luminance is near threshold, the threshold of the test field may be lowered (cf. Beitel, 1934, 1936; Karn, 1935; van den Brink & Bouman, 1954). When the inducing luminance exceeds a critical value, test threshold is raised (Beitel, 1936; Spencer, 1925). Both effects are largest when the test and inducing fields are adjacent and decrease as the distance between the stimuli is increased.

The purpose of the present experiment was to study the influence upon test threshold of the luminance, number, and location of inducers in the fovea and in the periphery. The inducers were a fixed distance from the test field while the distance between the inducers was varied.

METHOD

Apparatus.—A diagram of the optical system is shown in Fig. 1. The two sources (S) were 6-v. automotive lamps operated on direct current. Two lenses (C) in front of each lamp focused the filament on a flashed opal glass plate (G). The stop (P) in front of each opal glass made the focus at *S*'s eye 2 mm. in diameter, so that changes in pupil size did not alter the amount of light entering the eye.

A sector disk (D) continuously driven by a synchronous motor transmitted test flashes 5.3 msec. in duration. The *E* presented single flashes by opening a shutter (R) for one rotation of the disk. Lens L_1 collimated light from the opal glass. Test luminance was varied by Wratten neutral-tint filters (F) and a neutral-tint wedge and balancer (WB). A stop (T) at the principal focus of Lens L_3, formed a circular test field 26′ of visual angle in diameter. A semitransparent mirror (M) reflected the test beam and transmitted the inducing beam to Lens L_3, which brought both beams to a focus at *S*'s eye (E). Thus both test and inducing fields were presented in Maxwellian view.

The inducing light was collimated by Lens L_2. Its luminance was varied by Wratten neutral-tint filters (F). A stop (I) at the principal focus of Lens L_3 formed the inducing field. Each inducer was 10′ in diameter. The distance from the center of the test field to the center of each inducer was 25′. Thus the shortest distance from the circumference of the test field to the circumference of each inducer was 7′, and the diameter of the whole stimulus field was 60′. The stimulus field is shown in Fig. 2.

A luminous red fixation pattern at the same apparent distance as the stimulus field, was reflected from a microscope cover-slip between Lens L_3 and the eye. (This is not shown in Fig. 1.) When the fovea was stimulated, the fixation pattern was a line with a gap in it. The test field appeared in the center of the gap, which *S* fixated. For

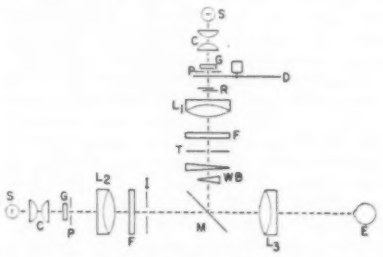

FIG. 1. The apparatus. Description in text.

[1] This article is based upon a dissertation submitted by the first author in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Faculty of Pure Science, Columbia University. The authors gratefully acknowledge the advice received from C. H. Graham, W. J. McGill, C. G. Mueller, and Yun Hsia of Columbia.

[2] Now with the System Development Corporation, Lodi, New Jersey.

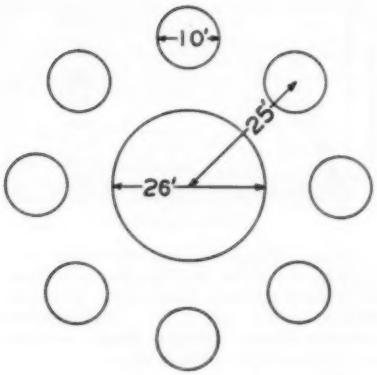

FIG. 2. The test field surrounded by eight inducers. Dimensions are in minutes of visual angle.

peripheral determinations the fixation pattern was a cross 10° temporal to the center of the test field.

The *S* sat in a dark room. A biting board and a head rest helped him to keep his right eye at the focus of the beams. He signaled his responses with a buzzer.

Procedure.—There were two male *Ss*. HR was emmetropic. IK wore glasses that corrected his myopia.

The *S* dark-adapted for 10 min. before foveal determinations and for 30 min. before peripheral determinations. The inducers were on continuously while the test field was flashed. The *S* light-adapted to the inducing luminance for 3 min. before thresholds were determined by the method of limits with the following inducers present: one, two adjacent to each other, two spaced on opposite sides of the test field, four adjacent, four evenly spaced around the test field, and all eight. Test luminance was altered by 0.05 log unit steps until *S*'s response changed from negative to positive or vice versa. Ascending and descending series were presented in ADDA order. The inducing luminance was varied from −2.70 to 3.51 log mL. in the fovea and from −4.49 to 3.73 log mL. in the periphery.

RESULTS

The results are plotted in Fig. 3–6. Each point is the average of 20 determinations, 10 for each of the two *Ss*.

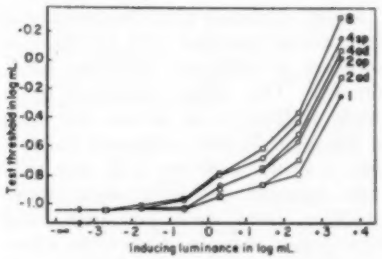

FIG. 3. Foveal test threshold as a function of inducing luminance for different numbers and locations of inducers.

The fovea.—Figure 3 shows test threshold as a function of inducing luminance for different numbers and locations of inducers. The inducers were not visible at luminances below −1.75 log mL. The luminance threshold of the test field was higher than that of the inducers because of the brief duration of the test exposure (cf. Graham & Margaria, 1935). There was no change in test threshold until the inducing luminance reached −0.62 log mL. At this value test threshold was raised slightly by four spaced inducers and by eight inducers. Test threshold rose for all numbers of inducers as the inducing luminance increased from 0.30 to 3.51 log mL. The rate of rise was steeper at higher luminances.

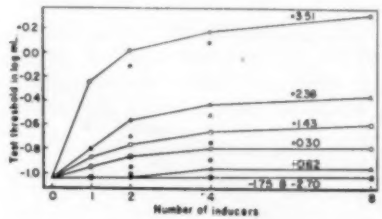

FIG. 4. Foveal test threshold as a function of number and location of inducers at different inducing luminances, which are labeled in log mL. The filled symbols represent adjacent inducers. Other conditions are represented by open symbols.

Figure 4 shows test threshold as a function of number and location of inducers at different inducing luminances. The effect produced by a single inducer was similar for each of the eight locations around the test field. From 0.30 to 3.51 log mL. two opposite inducers were more effective than two adjacent, and four spaced inducers were more effective than four adjacent. Test threshold rose as the number of inducers increased, but successive inducers produced decreasing increments in threshold log luminance.

The periphery.—Figure 5 shows test threshold as a function of inducing luminance for different numbers and locations of inducers. The inducers were not visible at −4.49 log mL. Test thresholds were slightly lower than with no inducers present, probably because additional dark-adaptation took place during the session. The inducers were dimly visible at −3.46 log mL., where test threshold began to rise. As inducing log luminance increased, the rate of rise in threshold log luminance first increased, then decreased, then increased again. The decrease in rate occurred at an inducing luminance of 0.50 log mL. for one inducer and for two adjacent inducers. For larger numbers of inducers, however, the decrease occurred at lower inducing luminances and higher test thresholds.

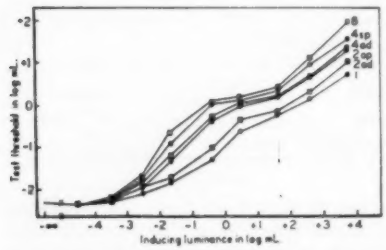

FIG. 5. Peripheral test threshold as a function of inducing luminance for different numbers and locations of inducers.

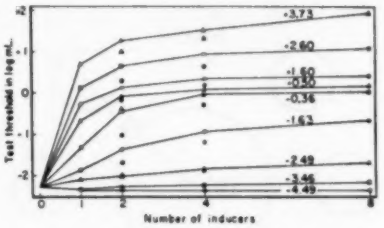

FIG. 6. Peripheral test threshold as a function of number and location of inducers at different inducing luminances, which are labeled in log mL. The filled symbols represent adjacent inducers. Other conditions are represented by open symbols.

The range of thresholds was larger for the periphery than for the fovea. Peripheral thresholds were initially below foveal thresholds, but rose above them as inducing luminance increased. With no inducers present the peripheral threshold was 1 log mL. less than the foveal. When the inducing luminance was about 0.5 log mL., peripheral thresholds became higher than foveal. At about 3.5 log mL. peripheral thresholds were 1 log mL. higher than foveal thresholds.

Figure 6 shows test threshold as a function of number and location of inducers at different inducing luminances. The effects were similar to, but larger than, those observed for the fovea. The rise in threshold produced by one inducer was about equal for each of the eight positions around the test field. From −1.63 to 3.73 log mL. two opposite inducers were more effective than two adjacent, and four spaced inducers were more effective than four adjacent. Increasing the number of inducers raised test threshold, but each additional inducer yielded a smaller increment in threshold log luminance.

TABLE 1

TEST THRESHOLDS IN LOG MILLILAMBERTS

Locus of Stimulation	Inducing Luminance (Log mL.)	Number and Location of Inducers							
		2 Inducers			4 Inducers			8 Inducers	
		Super.	Adj.	Opp.	Super.	Adj.	Spa.	Super.	Separ.
Fovea	+0.30	−0.93	−0.95	−0.87	−0.91	−0.89	−0.80	−0.89	−0.79
	+1.43	−0.85	−0.87	−0.77	−0.83	−0.75	−0.69	−0.80	−0.61
	+2.38	−0.65	−0.69	−0.57	−0.51	−0.52	−0.43	−0.36	−0.38
	+3.51	−0.10	−0.12	+0.01	+0.04	+0.07	+0.16	+0.19	+0.29
Periphery	−1.63	−1.71	−1.68	−1.36	−1.59	−1.18	−0.92	−1.46	−0.66
	−0.36	−1.07	−0.99	−0.42	−0.83	−0.27	0.00	−0.60	+0.05
	+0.50	−0.53	−0.36	−0.03	−0.43	−0.01	+0.11	−0.33	+0.19
	+1.60	−0.22	−0.14	+0.16	−0.19	+0.20	+0.35	−0.15	+0.42
	+2.60	+0.29	+0.30	+0.64	+0.45	+0.66	+0.97	+0.60	+1.09
	+3.73	+1.00	+1.02	+1.28	+1.28	+1.36	+1.58	+1.56	+1.94

Note.—Abbreviations: Super. = superimposed, Adj. = adjacent, Opp. = opposite, Spa. = spaced, Separ. = separate.

Distance effects.—Beitel (1936) has demonstrated that the rise in test threshold produced by an inducer, decreases as the distance between the test field and the inducer is increased. The present experiment shows that when the distance between the test field and a number of inducers is constant, the rise in threshold increases as the distance between the inducers is increased.

Test thresholds for a given number of inducers at different separations are presented in Table 1. The threshold values for superimposed inducers were obtained by interpolation from the curves for one inducer in Fig. 3 and 5 except for those in the last row of each table, which were obtained by extrapolation and so are less accurate. Superimposing two inducers is equivalent to increasing the luminance of one of them by 0.30 log units. Superimposing four or eight inducers is the same as increasing the luminance of one inducer by 0.60 or 0.90 log units, respectively.

From 0.30 to 3.51 log mL. in the fovea and from −1.63 to 3.73 log mL. in the periphery, two inducers raised test threshold more when they were opposite than when they were adjacent. Two adjacent inducers were about as effective as two superimposed. Four inducers generally raised threshold least when superimposed, more when adjacent, most when they were evenly spaced around the test field. It is consistent that four inducers were more effective when adjacent than when superimposed, while two inducers were about as effective adjacent as superimposed, because four adjacent inducers are spread over a larger distance than two adjacent. Finally, eight inducers surrounding the test field generally raised threshold more than eight superimposed inducers.

The "difference method" (Garrett, 1953, p. 227) was used to calculate t ratios for the data in Table 1. For the foveal results, the null hypothesis was rejected for the differences between two opposite and two adjacent and between four spaced and four adjacent, since the t ratios are 9.7 and 11.0 respectively, while $P(t > 5.8) = .01$ with 3 df. For the peripheral results, $P(t > 4.0) = .01$ with 5 df. The following differences were found to have t ratios larger than 4.0: Between

two opposite and two adjacent, $t = 7.9$; between four adjacent and four superimposed and between four spaced and four adjacent, the t ratios are 4.9 and 7.4, respectively; between eight separate and eight superimposed, $t = 9.7$.

Discussion

The results demonstrate that test threshold rises as inducing luminance increases, as the number of inducers increases, and as the distance increases between a given number of inducers that are a constant distance from the test field.

Karn (1935) and Beitel (1936) reported that test threshold was lowered by a threshold luminance inducer when both stimuli were flashed simultaneously. In the present experiment, however, test threshold was not lowered when the inducing luminance was near threshold, possibly because the inducers were on continuously while the test field was flashed.

In the fovea, test threshold rose as a function of inducing luminance (Fig. 3); the rate of rise was steeper at high inducing luminances. In the periphery the function was compound (Fig. 5): the rate of rise first increased, then decreased, then increased again. These results, like those for intensity discrimination and dark-adaptation, can be related to the duplicity theory of vision (cf. Hecht, 1937). The foveal function is attributed to cone activity. The lower branch of the peripheral function is attributed to rods; the upper branch, to cones. As the number of inducers is increased, the transition from rod to cone dominance appears to occur at lower inducing luminances and higher test thresholds. Under corresponding conditions peripheral cone thresholds are higher than foveal cone thresholds.

Eye movement, entoptic stray light, and neural interaction in the visual system may contribute to the rise in test threshold. Involuntary eye movements during fixation cause the image of the inducers to move on the retina. Riggs, Armington, and Ratliff (1954) observed that 1-sec. exposures permitted a retinal image to move an average of 3′ of arc. In the present experiment the inducers were exposed continuously, and there was a 7′-space between them and the test field. Thus the image of the inducers might occasionally fall on and light-adapt the retinal area stimulated by the test field at some later moment.

Stray light from the inducers that illuminates the retinal test area may increase the test luminance required for intensity discrimination. Stray light alone, however, cannot account for the effects produced by varying the number and location of inducers. If the amount of stray light on the retinal test area determined these effects, then a given inducing flux would produce a constant effect, regardless of the number of inducers that contributed it. Increasing the number of inducers from one to two would raise test threshold as much as doubling the luminance of one inducer, and adjacent inducers would be as effective as spaced inducers. But, at low inducing luminances, adding a second inducer opposite the first was as effective as a tenfold increase in the luminance of one inducer, while adding a second inducer adjacent to the first had little effect. In general, different numbers and locations of inducers that yielded the same flux of stray light over the test image produced different effects.

It is possible, however, that test threshold is influenced by the distribution as well as by the amount of stray light. Stray retinal illumination decreases with distance from the image (Boynton, Enoch, & Bush, 1954; DeMott & Boynton, 1958). Thus an inducer probably casts a gradient of stray light over the retinal test area. A more nearly uniform distribution of stray light can be achieved by reducing the size of the test field, so that a smaller part of the gradient covers the test image. If stray light alone raises threshold, then with a smaller test field different spatial arrangements that yield the same total flux should produce effects that are more nearly equal. Accordingly, a control experiment was performed using

a test field 3′ in diameter. Varying the number and location of inducers produced effects similar to those obtained with the 26′-test field. This result suggests that neural interaction contributes to the effects.

The effects of varying the distance between inducers resemble interaction effects recorded from the eye of Limulus by Hartline and Ratliff (1955). These authors found that the frequency of discharge from an illuminated receptor could be depressed by illumination of neighboring areas. With two inducing spots, the inhibitory effect was greater when the spots were on opposite sides of the test receptor than when they were grouped together on the same side. In view of the differences in method and organism between the work on Limulus and the present experiment, it cannot be asserted to what extent the resemblance reflects similar underlying processes.

Summary

A circular test field was surrounded by several spots of light (inducers), which were equidistant from the test field. The luminance threshold of the test field was determined as a function of the luminance, number, and location of inducers in the fovea and in the periphery.

1. Test threshold rose as inducing luminance increased. When luminance was expressed in logarithmic units, in the fovea the rate of rise increased as the inducing luminance increased, and in the periphery the rate of rise first increased, then decreased, then increased again.

2. Test threshold rose as the number of inducers increased, but successive inducers produced decreasing increments in threshold log luminance.

3. The rise in test threshold produced by a given number of inducers, increased as the distance between the inducers was increased. Thus, two or four inducers raised threshold more when they were evenly spaced around the test field than when they were grouped together on one side of the test field.

4. Varying the luminance, number, or location of inducers produced larger effects in the periphery than in the fovea.

References

Beitel, R. J. Spatial summation of subliminal stimuli in the retina of the human eye. *J. gen. Psychol.*, 1934, **10**, 311–327.

Beitel, R. J. Inhibition of threshold excitation in the human eye. *J. gen. Psychol.*, 1936, **14**, 31–61.

Boynton, R. M., Enoch, J. M., & Bush, W. R. Physical measures of stray light in excised eyes. *J. Opt. Soc. Amer.*, 1954, **44**, 879–886.

DeMott, D. W., & Boynton, R. M. Retinal distribution of entoptic stray light. *J. Opt. Soc. Amer.*, 1958, **48**, 13–22.

Garrett, H. E. *Statistics in psychology and education*. New York: Longmans, Green, 1953.

Graham, C. H., & Margaria, R. Area and the intensity-time relation in the peripheral retina. *Amer. J. Physiol.*, 1935, **113**, 229–305.

Hartline, H. K., & Ratliff, F. The organization of neural activity in the eye. *Symposium of physiological psychology, ONR sympos. Rep. ACR-1*, 1955, 92–101.

Hecht, S. Rods, cones, and the chemical basis of vision. *Physiol. Rev.*, 1937, **17**, 239–290.

Karn, H. W. The function of intensity in the spatial summation of subliminal stimuli in the retina. *J. gen. Psychol.*, 1935, **12**, 95–107.

Riggs, L. A., Armington, J. C., & Ratliff, F. Motions of the retinal image during fixation. *J. Opt. Soc. Amer.*, 1954, **44**, 315–321.

Spencer, L. T. The validity of Heymans' law. *Amer. J. Psychol.*, 1925, **36**, 427–433.

van den Brink, G., & Bouman, M. A. Variation of integrative actions in the retinal system: An adaptational phenomenon. *J. Opt. Soc. Amer.*, 1954, **44**, 616–620.

(Received October 8, 1959)

Journal of Experimental Psychology
1960, Vol. 60, No. 5, 290–298

ASSOCIATION AMONG STIMULI AND THE LEARNING OF VERBAL CONCEPT LISTS[1]

JACK RICHARDSON

Harpur College

One approach to the problem of the relationship of the areas of rote learning and concept learning is to manipulate the same variables within each area. If the variables and their effects are the same for both areas then there is no basis for inferring different processes. Environmental, task, and *S* variables may be manipulated with this approach but conclusions about the commonality of the learning process depends upon an adequate sampling of the tasks from the two areas. Since there is little agreement as to which tasks require concept learning a satisfactory sample of concept tasks is difficult to obtain.

The approach of the present study is somewhat different from the first in that it focuses upon task variables and is more concerned with the dimensions which are related to the differences in the difficulty of learning selected verbal lists than in the proper classification of the lists. The verbal lists were constructed from material selected from three different sources and were studied in an attempt to obtain information concerning the dimensions related to differences in learning. It is assumed that any information concerning relevant dimensions may be useful in analyzing other lists but it is not necessary to assume that the lists are representative of any particular class of materials. Thus the words *rote* and *concept* are used merely for identification and are not intended to be explanatory or to imply different processes.

The verbal lists used in this study differed only in the materials used as stimuli and the choice of the materials was dictated by the following considerations: (*a*) the materials all consist of common words so they permit the same kind of manipulation and presentation, (*b*) they are from sources which have rated or characterized words in different ways and some have been used for rote learning while others have been used for concept learning, and (*c*) it was expected that lists constructed from these materials would differ in difficulty.

In the previous study (Richardson, 1958) the number of adjectives used as stimuli in paired-associate lists was held constant while the similarity of the stimuli and the number of different responses were varied. It was found that high similarity among stimuli which required the same response facilitated learning while high similarity among stimuli which required different responses increased the difficulty of learning. One of the nine lists was constructed so that the same response was correct for each group of highly similar adjectives. This list was the least difficult. In view of this effect of the variation of stimulus similarity, Exp. I in the present study attempted to relate the learning and retention of three lists to ratings of association between stimuli. Haagen (1949) has demonstrated that ratings of similarity and

[1] This research was supported by a National Science Foundation grant to the Research Foundation of State University of New York. Appreciation is due Alice Kaminisky for cooperation which permitted completion of the association ratings.

ratings of association are highly correlated and it was assumed that instructions to rate association would be less ambiguous than instructions to rate similarity. The ratings of association were not directly related to the difficulty of learning the lists in Exp. I. Experiments II and III were attempts to obtain information concerning the basis of the differential difficulty of the lists.

METHOD

Material.—Four sets of 20 words were selected from three sources and each set consisted of four groups of four stimulus words and a name for each group. The words and sources of material are given in Table 1. The groups of low-similarity and high-similarity adjectives were selected from Haagen (1949) and the names are merely the fifth word from the same similarity group. The high dominance concepts are from Underwood and Richardson (1956) and the names are the descriptive adjectives which were given as responses to the nouns. The class concepts were selected from Reed (1946a) and the names are the classes to which the nouns belong.

Ratings.—Two different pairings of words were rated by two different groups of *Ss*. The first group of pairs consisted of the possible combinations of words within a group. The five words within each group in Table 1 were paired in all possible combinations so there were 10 pairs to each group and 40 pairs for each set of material. In addition to the 160 pairs formed from the material in Table 1 additional pairs were constructed in the same manner from two groups of medium dominance and two groups of low dominance words from Underwood and Richardson (1956). The 200 pairs of words were made into booklets by randomly assigning the pairs to position down the left of the page with each pair followed by a 15-point scale. The pairs were rated by 55 college students who did not serve in any learning experiment. The *S* was instructed to ". . . rate each pair according to the degree you think they are associated; that is according to the immediacy, compellingness, or consistency with which one word calls the other to mind." The low end of the scale indicated low association.

The second group of pairs consisted of words from different groups within a set of materials. Each of the four stimulus words

TABLE 1

LISTS AND SOURCES OF MATERIAL

Haagen (List L)		Haagen (List H)		Underwood & Richardson (List U)		Reed (List R)	
Low Sim. Adjectives	Name	High Sim. Adjectives	Name	High Dom. Concepts	Name	Class Concepts	Name
Artless Childlike Open Untaught	Sincere (BIH)	Double Dual Paired Twofold	Duplex (BIH)	Village Minnow Crumb Germ	Small (BIH)	Monkey Tiger Sheep Horse	Animal (BIH)
Heedless Fitful Fickle Giddy	Flighty (VOM)	Sacred Divine Hallowed Holy	Revered (VOM)	Barrel Doughnut Knob Balloon	Round (VOM)	Beet Potato Turnip Cabbage	Vegetable (VOM)
Faulty Blemished Injured Wanting	Unfit (RUK)	Spoken Oral Verbal Vocal	Voiced (RUK)	Bed Chamois Fur Velvet	Soft (RUK)	Red Green Blue Yellow	Color (RUK)
Fruitful Potent Pregnant Teeming	Profuse (ZET)	Tranquil Peaceful Quiet Serene	Placid (ZET)	Ether Garbage Gardenia Manure	Smelly (ZET)	Pine Maple Walnut Elm	Tree (ZET)

TABLE 2

MEAN ASSOCIATION RATINGS BETWEEN PAIRS OF WORDS

Material	Stimuli Within Groups	Stimuli and Names	Stimuli of Different Groups
Low Similarity (List L)	6.62	7.48	—
High Similarity (List H)	12.88	12.41	5.51
High Dominance (List U)	4.96	10.74	3.59
Class (List R)	9.21	12.50	4.19

in a group was paired in all possible combinations with the stimulus words from the other groups within the set of materials. These combinations were constructed from the high similarity, high dominance, and class materials in Table 1. Since there were 96 pairs for each set this gave 288 pairs of words and they were rated by 47 students at Broome Technical Community College. The procedure and instructions were the same as for the first group of ratings.

The mean association rating for each set of material is presented in Table 2. The first column of ratings is the association between stimuli within a group and the second column the association between the stimuli and the name. The last column gives the association between stimuli in different groups. In the learning experiments a single nonsense syllable was the response for a group of four words so that Table 2 presents the mean association between stimuli which had the same response, between the stimuli and a name for the group, and between stimuli which had different responses.

Lists.—The main lists will be specified here and only the deviations given for each experiment. The lists consisted of 16 paired associates with words as stimuli and nonsense syllables from Glaze (1928) as responses. The letter designating each list is given at the head of the column in Table 1 and the four groups of four words which were used as stimuli are presented in the column. A single syllable was the response for all four words in a group and the same four syllables were used as responses in all lists. The responses are given in parentheses in Table 1.

Procedure.—The 16 pairs in each list were arranged in three random orders subject to the restriction that not more than two pairs from a single group could appear in succession and that each of the 16 pairs must appear within an order. The words and syllables were projected with a 16-mm. strip-film projector, *S* could not see either *E* or the projector, and each *S* was run individually. The stimulus of a pair was presented for 2 sec. followed by a 2-sec. presentation of the stimulus and response together. There was a 4-sec. intertrial interval between each presentation of the 16 pairs.

Learning was by the anticipation method, the syllables were spelled by *S*, and, with the exception of Exp. I, the usual paired-associate instructions were used. The *S*s were all college students, had not previously served in a learning experiment, and were assigned to a counterbalanced order of conditions as they appeared at the laboratory.

EXPERIMENT I

Method

Lists H, U, and R were each learned by 20 *S*s, 10 for each of two sets of instructions. The first set of instructions (A) was for paired-associate learning and did not state that the stimuli could be grouped in any way or that a response would be correct for more than one stimulus. The other set (B) was the same except the following was inserted:

> There are 16 different words or a different word for each pair but there are only four different nonsense syllables. Thus each syllable will be correct for four different words. You should try to discover what is alike about the four words attached to a single syllable.

At the first session *S* learned the appropriate list to a criterion of two successive perfect trials. He returned one week later for recall and relearning.

When groups of similar stimuli are paired with the same response, then presentation of a single stimulus-response pair should increase the probability of other stimuli in the group eliciting the same response and this intratrial transfer should vary with the type of material. If the responses are presented during recall, then performance on the recall trial could be due to the intratrial transfer as well as the degree of learning and the amount of retention. In an attempt to avoid this the responses were not presented during recall. Recall consisted of five trials during which the list was presented in exactly the same way as during learning except the responses were not presented and there was no indication to *S* whether a response was right or wrong. After the five

recall trials the responses were again presented and relearning was continued to a criterion of two successive perfect trials.

Results

The mean and median numbers of trials required to learn to the criterion of two successive perfect trials are given in Table 3. An analysis of variance gave $F < 1$ for both instructions and interaction and $F = 8.96$ for lists, but the within-groups variance was heterogeneous by Bartlett's test. A median test indicated that List R was significantly less difficult than List U ($P < .01$) but that the difference between Lists H and U and between Lists H and R was of borderline significance ($.05 < P < .20$). It is concluded that the lists differ in difficulty. The class concepts, List R, are the least difficult while the high dominance concepts, List U, are the most difficult. The high similarity adjectives, List H, are intermediate in difficulty. The order of difficulty is not that expected from the association ratings between stimuli. Higher association between stimuli with the same response should decrease the difficulty but the ratings in Table 2 show higher association between stimuli with the same response for List H than for List R while the difference in association between stimuli with different responses is small. If the association between stimuli was the variable determining the speed of learning then List H should have been learned faster than List R.

After learning and relearning each *S* filled out a questionnaire which gave the four responses and asked what was alike about the words which had been paired with each syllable. The 20 *S*s who learned List R, the class concepts, indicated the class consistently by *animals*, *vegetables*, *colors*, and *trees*. In three cases *foods* was given instead of *vegetables*, and in one case *wood* instead of *trees*. Only one *S* indicated that he did not know, and this was with reference to only one group. In contrast to this, 8 different *S*s made a total of 14 inappropriate classifications of groups for List H while 18 *S*s made a total of 51 inappropriate classifications for List U. Verbalization of the basis of the groupings was related to the speed of learning.

The mean number of correct responses on the five recall trials and the first two relearning trials are presented in Fig. 1. An analysis of variance was computed on the number of correct responses for Recall Trials 1 and 5 and Relearning Trial 1, but none of the *F*s approached significance. Excluding recall and criterion trials, the mean number of trials required to relearn was 1.00, 1.90, and .75 for Lists H, U, and R, respectively. The differences are in the same direction as in learning but

TABLE 3

TRIALS TO LEARN TO CRITERION: EXP. I

Instructions	List H			List U			List R		
	Mean	σ_M	Mdn.	Mean	σ_M	Mdn.	Mean	σ_M	Mdn.
A	16.1	3.19	13.0	20.7	1.71	21.0	11.5	1.57	10.5
B	15.1	1.84	15.0	20.6	3.19	19.5	10.6	1.33	10.0
Combined	15.6	1.80	14.0	20.6	1.76	20.5	11.0	1.00	10.2

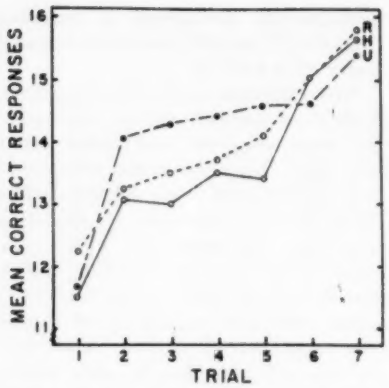

FIG. 1. The mean number of correct responses on the five recall trials and the first two relearning trials (Exp. I).

do not approach significance. The difference in the mean number of trials required to attain the criteria of one perfect trial and two successive perfect trials during learning was 2.70, 2.10, and 2.85 for Lists H, U, and R, respectively. This indicates that the rate of increase in performance was approximately equal at the end of learning so that differences in recall should not be a function of different degrees of learning resulting from different rates of learning.

As in the previous study (Richardson, 1958) there is an increase in performance over the five recall trials even though the responses were not presented during this time. Only three of the 60 *Ss* gave fewer correct responses on Recall Trial 5 than on Trial 1. For the combined lists there was 74% retention on Recall 1 which increased to 88% on Recall 5. The *Ss* had not participated in any other learning experiment and that may have been one reason for the high retention but there was also no control over rehearsal. There was forgetting from the last learning trial to the first recall trial so it was possible to have differences in retention. It seems that any differential rehearsal or differential effects of rehearsal would tend to increase the retention of the class concept list and it failed to show better retention. The number of correct responses during the recall trials increases as association between stimuli with the same response decreases but there is no indication that the differences in recall are reliable.

In attempting to relate the association between stimuli to the difficulty of learning the lists it was assumed that the difficulty would be determined by the stimulus relationships and that the association ratings would be an index of the transfer to be expected between stimuli. As the association between stimuli with the same response increased then the difficulty should have decreased, and as the association between stimuli with different responses increased the difficulty should have increased. This was not the case so it became necessary to re-examine the assumptions.

Even though the lists differed only in the words used as stimuli it is possible that the differential difficulty was due to the stimulus-response relationships rather than to the stimulus relationships, i.e., the differential difficulty may have been due to differences in the difficulty of the individual word-syllable pairs in the lists. If individual pairs in a paired-associate list do not interact in any fashion then increasing the length of the list by adding pairs of equal difficulty should have no effect on the number of repetitions required to learn, other than that produced by the increased interval between presentations of a pair. If learning one pair interferes equally with the learning of all other pairs then difficulty of a list should be an increasing, positively accelerated, function of the number of pairs in a list. However, if pairs are added to a list without increasing the number of different responses the results should depend upon the amount of transfer between stimuli with the same response as related to that between stimuli with different responses.

If the differential difficulty of the lists was due to the stimulus relationships then there should be an interaction

between type of list and number of stimuli, and the effect of increasing the number of stimuli should be related to the speed of learning in Exp. I.

Experiment II

Method

The lists in Exp. II differed only in the words used as stimuli and in the number of stimuli. Four lists were constructed with each of the four sets of material given in Table 1. One list contained four word-syllable pairs consisting of the first stimulus from each of the four groups listed in the columns in Table 1, the second list consisted of the first two stimulus words in each group paired with the same syllable, the third paired the first three stimuli with the syllable, and the fourth consisted of all four stimuli paired with the response. Thus the lists were of materials from Table 1 and consisted of either 1, 2, 3, or 4 stimuli for each of the four responses. The type of material will be designated by the letters given at the heads of the columns in Table 1 and the number of stimuli for each different response by a number following the letter. Thus List H-3 consisted of the first three stimuli in each group of high similarity adjectives, paired with the four nonsense syllable responses given in Table 1. There were 9, 6, 4, and 3 different orders of presentation for the lists with 1, 2, 3, and 4 stimuli per response, respectively, and each of the 16 lists was learned for 10 trials by a separate group of 12 *Ss*. A trial was one presentation of all pairs in a given list.

Results

An analysis of variance for correct responses during the 10 learning trials and for the correct responses on Trial 10 indicated that the type of list, stimuli per response, and interaction were highly significant but the error terms were not homogeneous. The mean correct responses on Trial 10 are given in Table 4. The number of correct responses increase as the number of stimuli increase and the amount of increase was directly related to the speed of learning the lists with four stimuli per response.

TABLE 4

Mean Number of Correct Responses on Trial 10: Exp. II

List	Number of Stimuli for Each Response			
	1	2	3	4
Low Sim. (L)	2.58	4.67	7.58	7.58
High Sim. (H)	2.67	5.83	7.00	12.17
High Dom. (U)	3.00	6.17	8.08	10.00
Class Concept (R)	3.42	5.92	10.67	14.50

In order to equate the possible number of correct responses for all lists the number of correct responses during learning was divided by the number of stimuli per response and the mean corrected learning scores are presented in Table 5 and the analysis of variance in Table 6. The *F*s are in the appropriate row for each term but the row for Within Groups contains the error term which was used to compute the *F*s in the column. The type of list and interaction is significant but the number of stimuli per response is not significant in the corrected score.

The different pairs did not appear equally often in the lists so part of the differential effect may be due to peculiarities of pairs within a group. As a final check the four pairs of items which made up the one stimulus

TABLE 5

Corrected Mean Number of Correct Responses during Learning: Exp. II

List	Number of Stimuli for Each Response			
	1	2	3	4
Low Sim. (L)	16.42	14.33	15.03	11.46
High Sim. (H)	14.83	16.33	17.36	21.62
High Dom. (U)	18.00	18.17	18.31	15.60
Class Concepts (R)	24.67	17.67	23.50	26.73

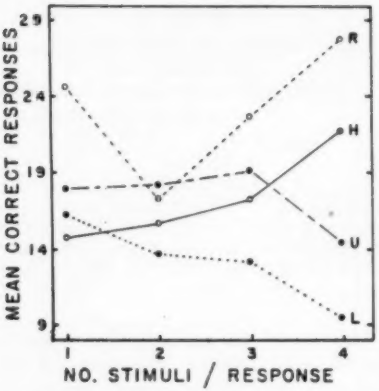

FIG. 2. The mean number of times the four pairs were given correctly during the learning of each list (Exp. II).

per response lists were traced through each list and the means of the number of correct responses during the 10 learning trials for only these four pairs are presented in Fig. 2 and the analysis of variance in Table 6. The mean number of correct responses for the four pairs correspond very closely to the corrected means presented in Table 4. The type of material and the interaction are significant but the number of stimuli per response is not a significant variable. Since these are the same four pairs, within each type of list, the interaction cannot be the result of differential difficulty of the pairs but must be the effect of the relation of the stimuli to other stimuli which are added to the list. If the difference between the correct responses on the list with one stimulus per response and the list with four stimuli per response, within each type of material, is evaluated by a *t* test, the high similarity lists (H) show a significant increase, the low similarity lists (L) show a significant decrease, and the changes in the concept lists are not significant. It seems that the lack of a significant increase in the class concept lists (R) is due to the R-1 list which shows significantly more correct responses than the other four-pair lists.

TABLE 6

*F*s AND ERROR TERMS FOR LEARNING: EXP. II

Source	*df*	Corrected Total Correct Responses	Total Correct: Four Pairs
List	3	14.22**	15.99**
No. Stimuli	3	1.09	1.01
Interaction	9	2.24*	2.98**
Within	176	45.48	49.92

* $P < .05$.
** $P < .01$.

In the lists with four stimuli for each different response the differential difficulty seemed to be due to differential transfer between stimuli but this was not directly related to the association ratings. The *S*s who had learned the lists in Exp. I rated the pairs of words with the same response and gave higher association ratings than *S*s who had not learned the lists. This may mean that, during the formal learning trials, there is a learning of relationships between stimuli and that this is a basis of transfer in addition to that based on the pre-experimentally learned relationships.

If the difficulty of the lists is due to intralist transfer based on pre-experimentally learned relationships between stimuli with the same response then the amount of transfer, as the result of learning the response for one stimulus of each group, should be inversely related to the difficulty of the lists. On the other hand, if the difficulty of the lists is due to the difficulty of learning the stimulus relationships during the formal learning trials, then the amount of transfer as the result of learning the name of the group should be directly related to the difficulty of learning the lists without the previous formal learning of the name.

Experiment III

Method

Two different conditions of first-list learning were used prior to learning each of the lists in Table 1. For each set of lists, a group of 12 *Ss* learned a four-pair list consisting of the first pair from each group in Table 1 and then learned the 16-pair list from the same set of materials, i.e., the second list consisted of the four pairs from the first list and 12 additional word-syllable pairs. A second group of 12 *Ss* learned a list consisting of the 16 stimuli with names as responses and then learned the list consisting of the same stimuli but with syllables as responses. In this case, the second list had the same stimuli but the responses were changed from the names to the syllables. There were eight groups of 12 *Ss* and the 10 trials of learning on the first list were followed immediately by 10 trials of learning on the second list.

Results

The *Ss* were assigned to lists in Exp. II and III at the same time so the learning of the transfer lists is comparable to the learning of the 16-pair lists in Exp. II and the first-list learning of the four-pair lists is that presented in Table 5. During the 10 trials of learning the name lists there were 63.0, 27.9, 15.8, and 17.7 mean errors for Lists L, H, U, and R, respectively. After Trial 2 there were no errors on List R and only two on List U. On the last trial there were 15.67 mean correct responses on List H and 13.58 mean correct responses on List L. Thus the learning was very rapid for Lists R and U, List H was somewhat more difficult, and Trial 10 performance on List L was still appreciably less than performance on the other lists.

The means of the number of correct responses during learning of the 16-pair lists are presented in Table 7. Analysis of variance on the total number of correct responses for the two transfer lists gave *F*s of 25.67, .72, and 3.76 for lists, first-list learning, and interaction, respectively, and the error term was 554.7. The amount of transfer for the lists is not in the same order as the difficulty of learning but there is an interaction between lists and conditions of first-list learning. If the amount of transfer due to learning one of the pairs from each group is compared with the association between stimuli with the same response (Table 2), it is apparent that the association is directly related to the amount of transfer from the four-pair lists. The amount of transfer after learning the name lists is inversely related to the association between stimuli.

TABLE 7

Mean Number of Correct Responses for First List Learning and for the Transfer Lists: Exp. III

List	Previous Learning		
	None	One Pair	Name
Low Sim. (L)	45.83	64.17	72.33
High Sim. (H)	86.50	113.17	99.58
High Dom. (U)	62.42	76.50	105.00
Class Concept (R)	106.92	129.17	122.42

Discussion

It has been demonstrated that the lists differ in difficulty and that the differences are the result of differential transfer between stimuli in the lists. Adding stimuli to List H and List R resulted in a decrease in the difficulty of the pairs of items and this agrees with the results of Reed (1946b) and Lloyd (1960) who used similar materials and required the same response for a group of stimuli.

The ratings of association between stimuli are not directly related to the difficulty of learning. List U is learned faster than List L and List R is learned faster than List H even though the association between stimuli would predict

the opposite. In both cases the less difficult list consists of groups of words which have a common verbal response. The groups in List U are nouns which have elicited the same response from *Ss* while the groups in List R are nouns commonly designated by the class to which they belong.

In addition to these descriptive differences in the lists there are several things which suggest that at least part of the transfer between stimuli with the same response is based on a mediating response. In Exp. I the number of verbalizations of the basis of classification of the groups of words was directly related to the speed of learning. In Exp. II the order of difficulty did not emerge until there were four stimuli for each different response, and this suggests that part of the transfer is based upon learning which occurs during the formal learning trials and which requires comparison of several stimuli. In Exp. III learning the name of the groups had a differential transfer effect on the lists.

The association between the stimuli and the names was not related to either the difficulty of the lists with syllables as responses or the difficulty of the lists with names as responses. Thus the association between the stimuli and the names fails to predict the difficulty of learning the names and, if the probability of a mediating response is the relevant dimension in these lists, it also fails to predict the probability of the occurrence of the mediating response in the experimental situation.

Summary

Paired-associate lists were constructed so that groups of four words were paired with the same nonsense syllable response. There were differences in the difficulty of learning the lists but the difficulty was not related to whether the lists were constructed from rote or concept materials and did not correspond to that expected from the ratings of association between the stimuli. There were no significant differences in recall of the lists one week after learning.

The number of different responses was held constant and the number of stimuli paired with each response was varied. There was an interaction between the lists and the number of stimuli which indicated that the differential difficulty was due to differential intragroup transfer between stimuli.

Differential amounts of transfer to the lists were obtained from learning one pair of each group of words and from learning the names of the groups. It was suggested the difficulty of the lists was determined by the difficulty of learning a mediating response for each group of stimuli and that this was not predictable from the association ratings.

References

Glaze, J. A. The association value of nonsense syllables. *J. genet. Psychol.*, 1928, **35**, 255–268.

Haagen, C. H. Synonymity, vividness, familiarity, and association value ratings of 400 pairs of common adjectives. *J. Psychol.*, 1949, **27**, 453–463.

Lloyd, K. E. Retention of responses to stimulus classes and to specific stimuli. *J. exp. Psychol.*, 1960, **59**, 54–59.

Reed, H. B. Factors influencing the learning and retention of concepts. I. The influence of set. *J. exp. Psychol.*, 1946, **36**, 71–87. (a)

Reed, H. B. The learning and retention of concepts: II. The influence of length of series. III. The origin of concepts. *J. exp. Psychol.*, 1946, **36**, 166–179. (b)

Richardson, J. The relationship of stimulus similarity and number of responses. *J. exp. Psychol.*, 1958, **56**, 478–484.

Underwood, B. J., & Richardson, J. Some verbal materials for the study of concept formation. *Psychol. Bull.*, 1956, **53**, 84–95.

(Received October 13, 1959)

Journal of Experimental Psychology
1950, Vol. 60, No. 5, 299-305

FIGURE-GROUND DOMINANCE AS A FUNCTION OF SECTOR ANGLE, BRIGHTNESS, HUE, AND ORIENTATION [1]

TADASU OYAMA

Hokkaido University

Many introductory textbooks describe Rubin's vase-profiles pattern and summarize the phenomenological characteristics and the determining factors of figure-ground perception. Few quantitative analyses have been done on this well-known phenomenon. The present research is an attempt to add some more quantitative data concerning the factors which make a given part of a visual field appear more frequently as "figure" and to obtain more precise knowledge of the effects of these factors. Area or the sector angle, brightness, hue, and orientation of the figure are the chief variables investigated.

General Method

Stimulus materials.—The stimulus patterns used in Exp. I, II, and III were similar to the drawing in Fig. 1. Each pattern consisted of a circle divided into six sectors, three alternate sectors of which have the same hue and brightness, and the other three have a different hue and brightness. In such a pattern the two alternating figures, which are shown at the bottom of Fig. 1, named the α-figure and the β-figure, are in Exp. IV, circular patterns with eight sectors were used.

The sectors were made from the standard colored papers for educational uses produced at Japan Color Research Institute and were pasted on cardboards. Each of these cardboards was presented behind the circular aperture of a vertical screen, 85 × 85 cm., standing on a table 115 cm. from *S*. The sector pattern was centered in the circular aperture, which was slightly smaller than the sector pattern and was on a level with *S*'s eyes. The screen consisted of white, black, or gray cardboard as determined by the experimental design. The size of the aperture was 16 cm. in diameter in Exp. I, II, and IV and 12 cm. in Exp. III. These corresponded to visual angles of 8° and 6°, respectively.

In Exp. I, II, and IV the stimulus pattern was illuminated by two 60-w. daylight bulbs which were placed above and behind *S*. In Exp. III, the bulbs were replaced by a Standard illuminator of Japan Color Research Institute which was matched with the CIE Standard Source C (6500° K), and was placed to the left of *S*'s head. In both conditions, the illuminance on the surface of the stimulus patterns was about 3 ftc. The power supply was 100-v., 50-cycle ac.

Recording apparatus.—Two telegraph keys were placed on the table in front of *S*. They were connected with two chronometers.

General procedure.—From 5 to 12 different patterns were presented to *S* in randomized order in one session. The *S* was instructed to continue to fixate binocularly the center of the stimulus pattern during the given observation period (30 sec. to 2 min.), and to press

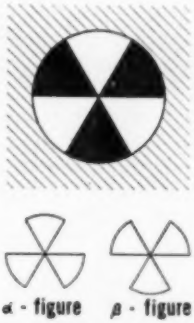

FIG. 1. An example of the figure-ground reversible patterns used in Exp. I, II, and III.

[1] A part of this study was carried out when the author was at the University of Tokyo. The author is indebted to Natsumi Hamamoto and Junichi Nakahara for their assistance in this research, and to Y. Takada, Hokkaido University, for suggestions improving the style of this paper.

one of the telegraph keys when the α-figure was seen and the other when the β-figure appeared. He was to press neither key when neither the α- or the β-figure was seen as figure, or in the absence of clear figure-ground perception. Rest periods of 1 to 3 min. were placed after every observation.

Measures.—The total time of pressing the α-figure key was designated T_α and that of β-figure T_β. The measure of relative dominance of the α-figure was computed by formula, $R_\alpha = 100\ T_\alpha/(T_\alpha + T_\beta)$. Similar formulae were used as the relative dominance of the white figure, Rw; that of the colored figure, Rc, etc.[2]

Subjects.—All *S*s were members of the laboratory staff or students of psychology. They all had normal color vision and normal acuity, although some of them wore glasses.

Experiment I: Sector Angle

Graham (1929) found that the set of sectors with the smaller angle in the six-sector pattern had a stronger tendency to be seen as "figure" than larger sectors and that the dominance of smaller sectors increased as the difference in angles increased. Goldhamer (1934), Oyama and Torii (1955), and Künnapas (1957) also studied this problem using similar patterns and obtained the same result. These results have been attributed to the effect of area by some *E*s and to the effect of proximity between the angled lines of contour of sector by others. No attempt was made to decide in the present experiment which was the determining factor, because it is difficult to vary these factors independently without the interference of such factors as surroundedness, good continuity, good gestalt, etc. Experiment I was designed to measure the effect of sector angle more systematically.

[2] In the experiments reported in this paper, the number of responses to a key varied from 5 to 10 per min. of observation; the time in which neither of the two keys was pressed varied from 0 to 15 sec. per minute.

Method

The α-sectors were made from black paper (2.4% in reflectance) and the β-sectors from white paper (88%) in B-series, and they were reversed in W-series. The surrounding field was medium gray (17%) in both series.[3] The angle of the α-sectors was varied as follows: 10°, 20°, 30°, 40°, 50°, 55°, 60°, 65°, 70°, 80°, 90°, 100°, and 110°; the angles of the β-sectors were, accordingly, 110°, 100°, 90°, 80°, 70°, 65°, 60°, 55°, 50°, 40°, 30°, 20°, and 10°.

Four *S*s observed each pattern twice. Each *S* served in eight sessions, and six or seven observations were obtained in each session. The observation time was always 60 sec. with rest intervals of 3 min.

Results

If the difference between black and white is neglected, the average relative dominance of α-sectors, R_α, as a function of angle, $\angle\alpha$, is shown by the solid line in Fig. 2. The dominance is highest when the angle is smallest, and decreases almost linearly as the angle increases. When the α-sectors have the same angle as the β-sectors, the dominance of α-figure is 53.4, and the curve is nearly symmetrical around this point. This fact

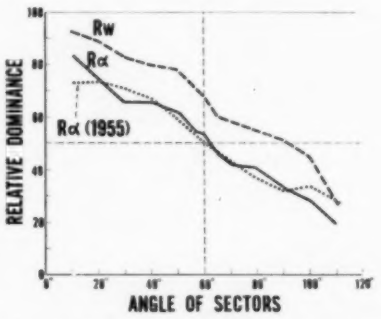

Fig. 2. The relative dominance of sectors as a function of sector angle.

[3] The reflectance values in this report are not those of samples of paper used but of standards from which the samples were copied.

suggests that there is little intrinsic difference in dominance between the α-figure and the β-figure. Extrapolation of the dominance curve to angles of 0° or 120° of the α-sector will not produce dominance of 100 or 0. There may be limits to the dominance regardless of the extremes of sector angle.

The curve of Fig. 2 is like the dotted curve which shows the results of Oyama and Torii (1955) who measured the dominance as a function of the sector angle in an outline pattern which had the same configuration as the stimulus patterns of the present study. But the latter flattens at both extremes, contrasting with the present curve which has nearly constant slope. There should be upper and lower asymptotic limits to dominance in the case of outline patterns. These limits, together with the abovementioned limits in the solid pattern, suggest that there is a constant term in dominance measure which is independent of the angle of the sectors, and that such a constant term is larger in outline patterns than in solid patterns. The constant term may represent the effect of the parts of the contour which remain constant even in the smallest sectors; this effect should be stronger in the outline pattern than in the solid pattern.

The broken line in Fig. 2 indicates the dominance of the white sectors, Rw, as a function of $\angle w$ when the effect of sector position is neglected. This curve is higher than the dominance curve of α-figure. When the white and black sectors have the same angle, the dominance of the white sectors is 17.6 higher than the neutral value, 50. The dominance becomes neutral when the angle of the white sectors is 90° and the angle of the black sectors is 30°. However, it is questionable that this dominance of the whiteness is independent of the brightness conditions of the surrounding field. The effect of brightness is described in the next experiment.

Experiment II: Brightness and Brightness Difference

Graham (1929) showed that the brighter figure was more dominant than the darker figure, but Goldhamer (1934) reported the reverse results. Harrower (1936) found that a white figure was dominant in a dark gray surrounding field, but a black figure became dominant in a light gray field. Brightness difference rather than the absolute brightness was the important factor of dominance. More recently, Oyama and Sasamoto (1957) found that the brighter field always dominated over the darker field when the difference between the logarithmic reflectance of the brighter field and that of the medium gray surrounding field was equal to the difference between the surrounding field and the darker field. These variables are re-examined in Exp. II.

Method

The angles of both α-sectors and β-sectors were kept constant at 60° throughout this experiment. The surrounding field was black cardboard, with a reflectance of about 8%, in the BS-series; and white cardboard (reflectance 80%), in the WS-series. In half of the stimulus patterns of the BS-series, α-sectors were always white (88%) and the brightness of β-sectors was varied in 10 steps from black to very light gray. In the other half, β-sectors were white and α-sectors were varied. In half of the WS-series, α-sectors were always black (2.4%) and β-sectors were varied in 10 steps from very dark gray to white, and these relations were reversed in the other half. The series of gray papers used in the stimulus patterns of these two series consisted of 11 papers, whose reflectances varied nearly logarithmically as follows:

2.4%, 3.8%, 5.8%, 8.5%, 12%, 17%, 24%, 33%, 46%, 63%, and 88%.

Five *Ss* served in four sessions. Each *S* observed each pattern of the two series once in randomized order during these sessions. Both observation time and rest intervals were 90 sec.

Results

Figure 3 shows the relative dominance of the constant white or black sectors, *Rw* or *Rb*, as a function of logarithmic reflectance of the variable brightness sectors. The solid line represents *Rw* in the BS-series, and the broken line *Rb* in the WS-series. The reflectance values of surrounding fields are indicated by short vertical lines. The broken line shows that the dominance of black sectors is greatest in combination with the brightest variable sectors and decreases as the brightness of the variable sectors decreases. The maximum of the solid line occurs with lower reflectance of the variable sectors. These facts mean that relative dominance of the white or black sectors is greatest when the variable sectors have nearly the same brightness as the surrounding field, and decreases as the brightness difference between the variable sectors and the surrounding field increases. Similar tendencies are observed in the dotted lines which represent the results of Oyama, Torii, and Hamamoto (1957) who studied this problem under similar conditions except that they used medium gray for the surrounding field. Thus, the results of the present and previous studies indicate that there is an effect of brightness difference, but that it is not so conspicuous as in the effect of sector angle in Exp. I or as the effect of hue in Exp. III.

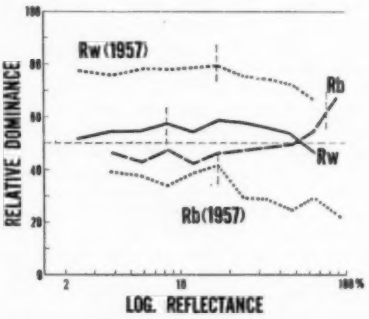

FIG. 3. The relative dominance of sectors as a function of brightness.

If the brightness difference were the sole factor involved in brightness variables, the black sectors, which have maximal brightness difference in the white surrounding field, should always dominate the variable sectors. However, this is not the case in the WS-series as shown by the broken line in Fig. 3. By similar reasoning, the black sectors in the medium gray surroundings were expected to dominate over the variable sectors under most conditions, but the lower dotted curve shows that this did not occur. To explain these facts, we must assume that black sectors are intrinsically less likely to be seen as "figure," or that the factor of absolute brightness as well as the factor of brightness difference is important in this perceptual function.

EXPERIMENT III: HUE

The effect of color on the figure-ground perception has already been discussed by Rubin (1921) who cited Ameseder's study which was concerned with the problem of attention rather than of figure-ground perception. Graham (1929) and Harrower (1936) also studied this problem but did not obtain clear-cut results.

In Exp. III, one set of the alternating sectors was a saturated color and the other set was a gray of the same brightness. Under these conditions, it was supposed that the effect of

brightness operated equally in both sets of alternating sectors, and if there was any difference between the dominances of the two sets, it should be attributed to the effect of hue.

Method

The surrounding field was black in the BS-series, and white in WS-series. The central part, 30 cm. wide and 20 cm. high, of the surrounding field, immediately adjacent to the sector pattern, was made from black (2.4%) or white paper (88%) and the remaining part from the black or white cardboards as in Exp. II.

The sector-patterns used in the two series were the same 24 patterns. The angles of both α-sectors and β-sectors were always 60°. In half of them, α-sectors were made from chromatic papers, β-sectors from achromatic papers; in the other half these relations were reversed. The chromatic papers were red (5.0R 4.0/12.0 in Munsell notation), orange (10.0R 5.5/11.0), yellow orange (7.5YR 7.0/11.5), yellow (7.5Y 8.0/11.0), yellow green (7.5GY 7.0/8.5), green (5.0G 5.5/8.0), blue green (7.5BG 5.0/9.0), greenish blue (2.5B 4.0/8.5), blue (2.5PB 4.0/10.0), blue purple (7.5PB 3.0/14.0), purple (5.0P 3.0/11.0), and red purple (2.5RP 3.5/11.0). The reflectance values of the achromatic papers paired with the above were 12%, 24%, 46%, 63%, 46%, 24%, 17%, 12%, 5.8%, 5.8%, and 8.5% respectively.

Five *Ss* served in four sessions, two for each of the two series. They observed each pattern once. The observation period was reduced to 30 sec., because longer period of observation resulted in chromatic adaptation. The rest period was 1 min.

Results

The solid line in Fig. 4 indicates the results of BS-series, the broken line those of WS-series.[4] The two curves are markedly similar. The dominance is highest at red and lowest at blue. Dominance was a decreasing function of the displacement of hue on both sides of red in the color circle.

[4] One of five *Ss* gave almost flat curves while the rest produced curves similar to the average curves.

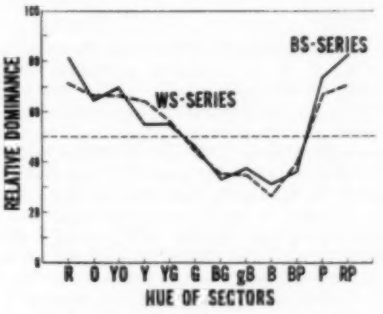

FIG. 4. The relative dominance of sectors as a function of hue.

This course of change of the dominance is very much like that of the effect of hue on apparent depth which was studied by Oyama (1958) and Oyama and Yamamura (1960) and we assume that there is some relationship between these two phenomena. Approach in depth corresponds to dominance, and the recession in depth corresponds to the weakening of the tendency to become "figure."

EXPERIMENT IV: ORIENTATION

Rubin (1921) found that a cross which consisted of vertical and horizontal arms had a stronger tendency to be seen as "figure" than a cross with oblique arms. This has been recently confirmed by Künnapas (1957) who studied this problem quantitatively using an eight-sector pattern.

Goldhamer (1934), Oyama and Torii (1955), and Oyama and Sasamoto (1957) found that there was little difference between the dominances of α- and β-sectors in the six-sector pattern. The results of Exp. I, II, and III in the present study showed practically the same tendency. Experiment IV was designed to analyze more systematically the effect of orientation in the eight-sector pattern.

Method

Eight kinds of eight-sector Patterns A, B, C, D, E, F, G, and H were used. The angle of the sectors was always 45°. In Pattern A, the vertical and horizontal sectors were made from white paper and the rest from black paper. Rotated clockwise 11.25°, Pattern A became Pattern B. Patterns C, D, E, F, G, and H were also the variations of Pattern A which were produced when Pattern A was rotated 22.5°, 33.75°, 45°, 56.25°, 67.5° and 78.75° respectively. The surrounding field was made from light gray cardboards whose reflectance was about 45%.

Four *Ss* observed each pattern twice in four sessions. The observation time was 90 sec., and the rest interval was 2 min.

Results

Figure 5 shows the dominance of white sectors as a function of the angle of inclination. The dominance of white sectors is greatest in Pattern A in which the white sectors extend in vertical and horizontal directions while the black sectors extend in oblique directions, and smallest in Pattern E in which the white sectors are in the diagonal position and the black sectors are in the vertical-horizontal position. This means that the sector-figure, either white or black, shows its maximal value of dominance in the vertical-horizontal position and its minimal value in the diagonal position. Between these extreme values, the dominance decreases and increases gradually as a function of angle of inclination.

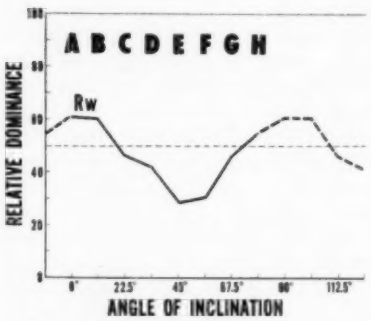

FIG. 5. The relative dominance of sectors as a function of orientation.

If we rotate Pattern A 90° or more, we merely reproduce Patterns A, B, C, etc. Consequently the curve of dominance as a function of inclination should be a cyclic one as shown by the broken line in Fig. 5, whose cycle is 90°. Similar cyclic curves have been found in many visual phenomena, geometrical illusions, figural aftereffects, etc. (Morinaga, 1933; Nozawa, 1949; Obonai, 1930; Oyama, 1960). In the vertical or horizontal directions, the principal axes of our visual space, stronger tendency to be seen as "figure," and smaller amount of illusion in Poggendorff's and Zöllner's figures, and smaller aftereffects of Gibson's curved line have been observed. They all represent the anisotropic character of our visual space.

SUMMARY

An attempt was made to analyze quantitatively the effects of the area or sector angle, brightness, hue, and orientation upon figure-ground perception. The figure-ground reversible pattern was a circle divided into six or eight sectors. The alternate three or four of the sectors had a constant hue and brightness, and the other set had a different hue and brightness. The percentage of the time the former set of sectors was seen as "figure" as compared to the total time either set of sectors was seen as "figure" in the given period of observation was employed as the criterion of relative dominance.

The experimental results showed:

1. The set of sectors with the smaller angle had a stronger tendency to be seen as "figure" than larger sectors. The dominance of smaller sectors increased almost linearly as the difference in angle increased.

2. The effect of brightness was not as conspicuous as the effect of area or sector angle. The dominance of sectors increased as the brightness difference between the sectors and the surrounding field increased. Besides this effect of brightness difference, there appeared to be an effect of absolute

brightness in that dark sectors were less frequently seen as "figure."

3. When the difference of brightness was eliminated, red sectors were the most dominant and blue sectors were the least. The dominance varied continuously as a function of hue between these extreme values.

4. The set of sectors which extended in vertical and horizontal directions was more dominant than the set of sectors which extended at oblique angles. The dominance varied as a cyclic function of the angle of inclination of the sectors.

REFERENCES

GOLDHAMER, H. The influence of area, position, and brightness in the visual perception of reversible configuration. *Amer. J. Psychol.*, 1934, **46,** 189–206.

GRAHAM, C. H. Area, color, and brightness difference in a reversible configuration. *J. gen. Psychol.*, 1929, **2,** 470–481.

HARROWER, M. R. Some factors determining figure-ground articulation. *Brit. J. Psychol.*, 1936, **26,** 407–424.

KÜNNAPAS, T. M. Experiments on figural dominance. *J. exp. Psychol.*, 1957, **53,** 31–39.

MORINAGA, S. Untersuchungen über die Zöllnersche Täuschung. *Jap. J. Psychol.*, 1933, **8,** (Japanese text, 195–242; German summary, 23–25).

NOZAWA, S. An experimental study on the after effect of curved line: II. Paper read at Japanese Psychol. Ass., 1949.

OBONAI, T. Experimentelle Untersuchungen über "das periodische Phänomen der Abweichung." *Jap. J. Psychol.*, 1930, **5,** (Japanese text, 469–474; German summary, 11–13).

OYAMA, T. A measurement of advancement and recession of colored surfaces. *J. illum. Engng. Inst. Japan*, 1958, **42,** 526–531. (Japanese text with English summary.)

OYAMA, T. Japanese studies on the so-called geometrical-optical illusions. *Psychologia*, 1960, **3,** 7–20.

OYAMA, T., & SASAMOTO, S. Experimental studies of figure-ground reversal: II. The effects of brightness, brightness gradient, and area in the tachistoscopic presentation. *Jap. J. Psychol.*, 1957, **28,** (Japanese text, 18–27; English summary, 64–65).

OYAMA, T., & TORII, S. Experimental studies of figure-ground reversal: I. The effects of area, voluntary control and prolonged observation in the continuous presentation. *Jap. J. Psychol.*, 1955, **26,** (Japanese text, 178–188; English summary, 217–218).

OYAMA, T., TORII, S., & HAMAMOTO, N. Experimental studies of figure-ground reversal: III. The influence of size, brightness gradient, illuminance, and the number of sectors upon the rate of reversal. *Jap. J. Psychol.*, 1957, **28,** (Japanese text, 210–222; English summary, 249–250).

OYAMA, T., & YAMAMURA, T. The effect of hue and brightness on the depth perception in normal and color-blind subjects. *Psychologia*, 1960, **3,** in press.

RUBIN, E. *Visuell wahrgenommene Figuren.* Copenhagen: Gyldendalske, 1921.

(Received October 16, 1959)

Journal of Experimental Psychology
1960, Vol. 60, No. 5, 306–309

EFFECTS OF RANGE OF PAYOFFS AS A VARIABLE IN RISK TAKING[1]

JEROME L. MYERS AND ERNEST SADLER

University of Massachusetts

This paper reports an experiment which investigates the decision to take risks, i.e., to gamble. On each of 100 trials *S* has the choice of an event with uncertain outcome and an event with known outcome. An optimal strategy exists and is a function of the arithmetic means of the two types of outcomes. Holding these means constant, the effect of varying the range of payoffs for risk taking, i.e., the range of uncertain outcomes, has been studied.

Edwards (1954) has previously investigated the role of variability of payoffs. His *S*s were required to choose between two bets having equal means but different variances. Although individual *S*s showed variance preferences, no strong trend existed over all *S*s. The present study differs from that of Edwards in that *S*s must choose between two events, one of which always has zero variance (the known outcome). Furthermore, our experimental approach permits an evaluation of the joint effect of the known payoff size and the range of unknown payoffs, as well as an evaluation of possible range preferences.

METHOD

Procedure.—All *S*s were exposed to two decks of 3 × 3 white cards. Deck 1 contained the known outcomes and consisted of 50 cards with +1 written on them and 50 cards with −1. Deck 2 contained the unknown outcomes, and its range of values was the major independent variable in the study. For Group R(4) (the range of absolute values was 4) Deck 2 consisted of 100 cards, 10 each of the following values: −6, −5, −4, −3, −2, +2, +3, +4, +5, and +6. For Group R (9) Deck 2 again consisted of 100 cards. Each value from −11 to +11 was represented five times with the exception of −1, 0, and +1 which were again absent. For Group R (14) Deck 2 contained only 90 cards, each value from −16 to +16 appearing three times, with the exception of −1, 0 and +1. Fifty poker chips were given to each *S* at the start of each session.

The procedure was the same for all three groups. At the beginning of each of 100 trials, *S* took the top card from Deck 1, which was face down. This card constituted the known payoff. Having seen this payoff *S* could either accept or reject it. If he accepted it, he gained or lost one chip depending on whether the card was +1 or −1. If *S* rejected the Deck 1 value, he was required to accept the top card from Deck 2, which was also face down. He then won or lost a number of chips corresponding to the number on the Deck 2 card.

At the start of the session *S*s were told that Deck 1 consisted of plus and minus ones and that Deck 2 consisted of larger positive and negative values. They were also told that the object of the experiment was to win as many chips as possible over the 100 trials. The *E* then shuffled both decks.

At the end of each session, *S*s answered an open-end questionnaire in which they were asked to describe their strategies, and to estimate the means of Decks 1 and 2 upon an 11-point scale, from −5 to +5.

Optimal strategy.—In order to maximize the expected value of the game, *S* should always accept a known payoff of +1, and reject a known payoff of −1. In other words, risks should always be taken when the alternative option is a loss of one point, and should never be taken when the alternative option is a gain of one point. This strategy is optimal when, as in this case, the means of Decks 1 and 2 are zero, regardless of the range of Deck 2 payoffs.

Subjects.—Sixty women enrolled in the elementary psychology course at the Uni-

[1] Support for this investigation was provided by the University of Massachusetts Teacher's Research Fund.

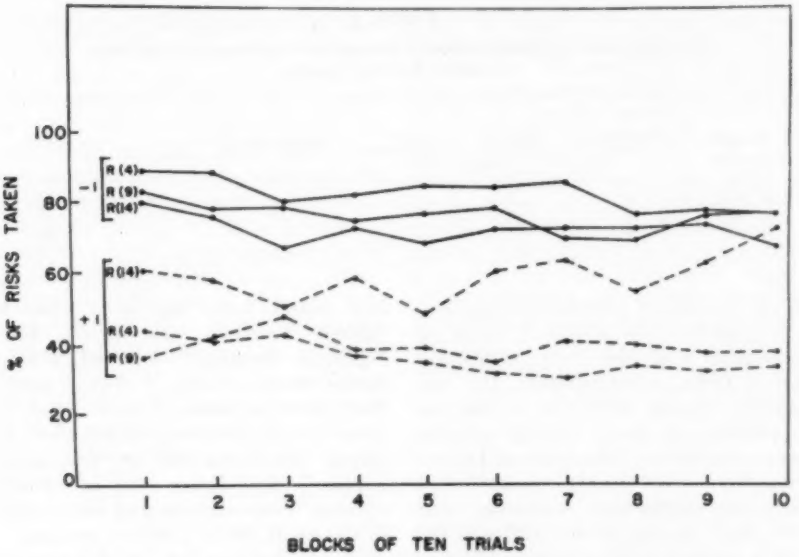

FIG. 1. Average percentage of risks taken by each group on +1 and on −1 trials, for each block of 10 trials.

versity of Massachusetts were randomly selected and assigned to the three groups.

RESULTS

Percentage choice.—Figure 1 presents the mean proportion of risk-taking responses for each block of 10 trials for each group, on both −1 and +1 trials. Thus, if there were six +1 and four −1 trials in a 10-trial block, and *S* gambled on two of the plus trials and on three of the minus trials, his scores for that block would be 33.3% and 75%, respectively. Figure 1 shows that, for all three groups, behavior was in the direction of the optimal solution. There were many more Deck 2 choices on −1 than on +1 trials. Group R (4) (narrowest range) most closely approximated the optimal solution, Group R (9) was next, and Group R (14), though generally gambling more often on −1 than on +1 trials, deviated most from the optimal bounds of 0 and 100%.

The results of an analysis of variance of the $\sin^{-1}\sqrt{\%}$ are shown in Table 1. "Value" refers to the comparison of +1 and −1 trials, and

TABLE 1

ANALYSIS OF VARIANCE OF PERCENTAGE OF RISKS TAKEN

Source	*df*	*MS*	*F*
Range (R)	2	4,554.49	1.07
Ss/R[a]	57	4,255.62	
Value (V)	1	223,335.63	75.79
Trials (T)	9	399.60	1.23
R × V	2	22,481.46	7.63
R × T	18	363.45	1.11
V × T	9	240.28	.63
R × T × V	18	274.07	.72
Ss × V/R[b]	57	2,946.64	
Ss × T/R[c]	513	326.12	
Ss × V × T/R[d]	513	378.59	

[a] Error term for R.
[b] Error term for V, R × V.
[c] Error term for T, R × T.
[d] Error term for V × T, R × V × T.

TABLE 2

CLASSIFICATION OF QUESTIONNAIRE RESPONSES AND AVERAGE ESTIMATES OF DECK 1 AND 2 MEANS

Group	Gambler's Fallacy	Optimal Strategy	No Solution	Miscellaneous	Deck	
					1	2
R (4)	0	3	7	2	−.18	.45
R (9)	2	9	6	5	.25	.85
R (14)	9	0	11	4	.05	1.20

is a significant source of variance ($P < .001$). The Range × Value interaction was also very significant ($P < .001$), substantiating the impression gained from Fig. 1 that the approximation to the optimal solution was a function of the range of Deck 2 payoffs. The Range main effect was not significant, indicating that the three groups did not differ in the total number of risk-taking responses. It necessarily follows that increases in the range raised the +1 curve about the same amount as the −1 curve was lowered. Neither the Blocks main effect nor any interactions involving Blocks were significant. This suggests that the effect of Value and the joint effect of Value × Range took place by the end of the first 10 trials. This inference is rather tentative since the *Ss* × B/R error term may be inflated owing to the shuffling of the decks for each *S*.

Questionnaires.—Table 2 presents a classification of *Ss*' responses to the questionnaire. The questionnaire was introduced into the procedure after nine *Ss* were run, all in Group R (4). Since some *Ss* had used more than one strategy during the 100 trials, they were represented in more than one cell. Thus the row totals are slightly larger than the number of *Ss* in each group.

The "gambler's fallacy" refers to such responses as "if more than two pluses came up in a row, I figured a minus was due." The "optimal strategy" covered statements ranging from "I always gambled after a minus, never after a plus" to "I always gambled after a minus, and about half the time after a plus." The latter type of "semi-optimal" response occurred frequently in Group R (9).[2] The "no strategy" class consisted mainly of statements such as "I guessed." The "miscellaneous" category included all other statements, mainly hypotheses about patterns of payoff. It appears that only Group R (14) was influenced by the gambler's fallacy. Furthermore, no one in this group verbalized any tendency to gamble more on −1 than on +1 trials, although it is evident in Fig. 1 that such was their behavior.

Table 2 also presents the average estimates of the deck means by each group. All estimates of the Deck 1 mean are close to zero. The estimates of the Deck 2 means are less accurate, and there is a tendency to overestimate, which increases with increases in range. Differences in the shapes and variances of the distributions precluded the use of

[2] The true percentage of those who consciously used the optimal solution in Group R (4) is probably higher than Table 2 indicates since most of the nine *Ss* who were not questioned actually performed somewhat better than those questioned.

analysis of variance to compare groups and the use of most nonparametric tests was ruled out by the numerous ties in the data. Therefore the data for each deck were classified in a 3×3 contingency table, the rows representing the groups, the columns representing ratings from -5 to -2, -1 to $+1$, and $+2$ to $+5$, respectively. Chi square analyses were applied to the data for each deck, yielding results which did not approach significance. Nevertheless, the observed bias in Deck 2 estimates is consistent with the results of Irwin and Smith (1957) who found that as the variance increased, *S*s required more information to estimate the mean of a deck.

Discussion

Apparently, a significant difference in risk-taking behavior on $+1$ and -1 trials exists by the end of 10 trials. Since this difference exists by the end of 10 trials, and since the distance between the $+1$ and -1 curves is rather great, it seems reasonable to conclude that many *S*s enter the experiment with a tendency to gamble more on -1 trials. The lack of a significant Range main effect indicates that there are no significant differences in the total number of gambles taken as a function of the range of the gamble. This substantiates Edwards' finding that variance preferences are of secondary importance in gambling situations.

Increased deviations from optimal solution with wider ranges of Deck 2 payoffs can be explained if the following assumption is made. Assume that as the range of payoffs increases, response tendencies are based upon an increased number of factors, which are unrelated to the means of the deck and to the particular Deck 1 value present on a given trial. Consider, for example, the gambler's fallacy. It is a noticeable factor in the behavior of the widest range group only. This is reasonable since the large positive and negative payoffs might well command more attention to the sequence of payoffs. Thus *S*'s response on a given trial is a function of whether the previous trials have yielded a sequence of pluses or a sequence of minuses. If the sequence effect is compatible with the optimal strategy (e.g., if there has been a sequence of negative payoffs in Deck 2 and the Deck 1 card on a given trial is a -1) the risk-taking curve might be moved slightly towards its optimal boundary. If the sequence effect and the Deck 1 card are not compatible, the curve will move away from its optimal bound. There is more room for performance to deviate from an optimal solution than to approach it. Thus, the net result is that the curves move away from the 0 and 100% bounds, and tend to converge, as the range increases. At least two other factors, besides the gambler's fallacy, may become more important as the range of payoffs increases, and should affect behavior in the manner suggested for the gambler's fallacy. The first of these is the reinforcement or punishment effect resulting from the most recent gamble; the second is the trial-to-trial fluctuation of *S*'s pile of chips.

Summary

Human *S*s were required to choose between a known payoff and an unknown payoff (the risk) on each of 100 trials. As the range of unknown payoffs increased, *S*s deviated more from an optimal strategy which required that they always gamble when the known payoff was negative, and never when it was positive. All *S*s did gamble more when the known payoff was negative.

References

Edwards, W. Variance preferences in gambling. *Amer. J. Psychol.*, 1954, **67**, 441–452.

Irwin, F. W., & Smith, W. A. S. Value, cost, and information as determiners of decision. *J. exp. Psychol.*, 1957, **54**, 229–232.

(Received October 16, 1959)

Journal of Experimental Psychology
1960, Vol. 60, No. 5, 310–313

EYELID CONDITIONED RESPONSES WITH VARIOUS LEVELS OF ANXIETY[1]

MARTIN R. BARON AND JAMES P. CONNOR[2]

Kent State University

Numerous studies have been reported in which the influence of drive level upon classical eyelid conditioning has been investigated. Two types of studies have been performed to test implications of the assumption that the drive level, *D*, in aversive situations is a function of the strength or magnitude of an emotional response aroused by aversive stimulation. In one, the intensity of the aversive stimulus has been varied; in the other, groups differing in emotional responsiveness have been conditioned and their performance compared (Spence, 1956, 1958). In keeping with predictions made on the basis of Hull's (1943) postulated relationship between *D* and habit, *H*, intensity of the UCS (Spence, 1953; Spence & Taylor, 1951), anxiety (Spence & Beecroft, 1954; Spence & Farber, 1953; Spence & Taylor, 1951, 1953; Taylor, 1951) and, more recently, physiological measures of emotionality (Runquist & Ross, 1959; Runquist & Spence, 1959) have been found to be related to eyelid CR performance.

Hull's formulation predicts that increases in drive level will result in a monotonic increase in eyelid CR performance. While most studies dealing with anxiety level as measured by the Taylor scale (1953) have compared extreme MAS *A*-score groups, Spence and Taylor (1953) present evidence showing the overall relationship to be linear. Since the authors mention that this finding of linearity differs from that suggested in unpublished studies, and since evidence of a nonmonotonic relationship between MAS *A* scores and performance on high association verbal responses had been found (Baron, Billey, & Connor, 1958), it appeared desirable to check further the shape of the function relating MAS *A* scores and eyelid CR performance.

The purpose of this experiment was to repeat, in functional characteristics, the eyelid conditioning study performed by Taylor (1951) employing *Ss* at various levels of anxiety as measured by the MAS.

METHOD

Subjects.—Fifty *Ss* were selected so as to be representative of the MAS *A*-scale distribution of scores of 131 students in general and educational psychology classes at Kent State University. Students obtaining MAS *L* scores of 7 or above were eliminated and quintiles computed. Five male and five female *Ss* were selected from each interquintile range, told they had been randomly chosen, and asked to take part in the experiment. The data from two *Ss* were not included in the results; one male *S* from the anxious end of the scale met the criterion for a "voluntary blinker," and one female from the nonanxious end of the scale terminated the conditioning session when she became ill.

[1] This report is based upon a thesis submitted by the second author to the Graduate School of Kent State University in partial fulfillment of requirements for the MA degree. An abbreviated form of this report was presented at the APA meetings, 1958, in Washington, D. C.

[2] Now with Psychological Services, Stark County Public Schools.

Apparatus.—The experiment was carried out in a sound-deadened experimental room. To seat the subject, an adjustable piano stool was used. Adjustable head and chin rests were employed to keep the position of *S*'s head fixed.

The CS was a light presented through a 6-in.-square milk glass window; it consisted of an increase in brightness from 6.5 to 8.6 ft-L as measured by a Luckiesh-Taylor GE brightness meter. The UCS, a puff of air, was administered to the cornea of the right eye; it was produced by a column of mercury, magnetically released, falling 80 mm. in a manometer $\frac{3}{4}$ in. in diameter. The size of the opening in the nozzle of the tube was $\frac{1}{16}$ in.

Presentation of stimuli was controlled by the fall of a mechanical pendulum, magnetically released, which opened and closed mercury switches.

The eyeblink was recorded by a photoelectric cell which was activated when a beam of light, projected through a 17.5-cm. focal length lens from a 32-cp automobile headlight, was cut by the closure of the eye with an artificial aluminum-foil eyelash which was attached to *S*'s eyelid. The blink, affecting the photoelectric cell, indirectly stopped an electric clock (started with the onset of CS) thus providing an objective measure of the latency of occurrence of the blink. A red light, also activated by the photoelectric cell, signaled *E* when *S* had his eyes closed.

Conditioning procedure.—Conditioning was carried out in one session with the average time for each *S* approximating 1 hr. After the artificial eyelash was fixed into position, *S* was seated and instructed to fixate the milk glass window located slightly above eye level on the wall of a booth 18 in. from his eyes.

Further instructions, designed neither to inhibit nor facilitate eye blinks, were read. Following this, each *S* received three presentations of the CS alone and two presentations of the UCS alone in the following order: CS, CS, UCS, CS, UCS; intervals between presentations were 15, 20, or 25 sec. randomly varied. All *S*s then received 90 paired presentations of CS and UCS. Employing criteria described by Goodrich, Ross, and Wagner (1957), *S*s who blinked on 5 or more of the first 10 trials with a latency of less than .30 sec. were considered voluntary blinkers and their data discarded.

While many of the above characteristics of apparatus, stimuli, and method of recording are different in some respects from those employed by Taylor (1951), the time relationships in the two studies are identical. The duration of the CS on each trial was .55 sec.; the UCS was presented simultaneously with the CS offset; and intertrial intervals were of 15, 20, and 25 sec., averaging 20 sec. and fixed according to a prearranged schedule. It should be noted, however, that while Taylor (1951) employed a ready signal on each trial, in accordance with the conditioning procedure commonly followed in the Iowa laboratories, no ready signal was given in this experiment. A response was considered a CR whenever the measured latency was between .20 and .50 sec.

Results and Discussion

Despite some differences in procedure and method of recording, frequency of CR for the two extreme MAS *A*-groups in this experiment are compared in Table 1 with data obtained by Taylor (1951) from "comparable" *S*s. Here it can be seen that the CR frequencies, as measured by Taylor (1951), were slightly higher than those of the two extreme groups in the present investigation. Neither the difference between the two anxious groups nor between the two nonanxious groups was found to be significant.

The conditioning curves over 80 trials, presented in Fig. 1, are seen to be higher with higher levels of MAS *A* scores. The total number of CRs per *S* was correlated with

TABLE 1

Mean Percentage of Eyelid CRs by Anxious and Nonanxious *S*s over 80 Trials in Taylor's Study (1951) and This Investigation

Group	Taylor's Study	This Investigation
Anxious	59.75	47.90
Nonanxious	29.14	27.47

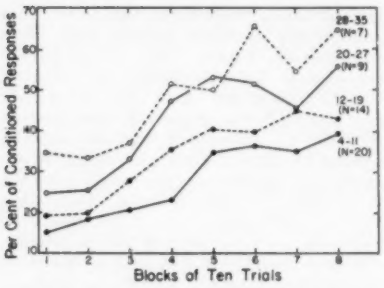

FIG. 1. Percentage of CRs in each block of 10 trials for four groups of *Ss* at different MAS *A*-score levels.

A scores; $r = .40$. This value is reliable ($P = .01$), and the relationship does not differ significantly from a straight line. This finding is consistent with predictions based upon the assumption of a multiplicative relationship between *D* and *H*. It should be noted that while no verbal ready signal was employed in the present study, and was included in the conditioning procedure in the Spence and Taylor (1953) investigation, the shape of the function was the same.

From inspection, the curves appear to be essentially parallel, with crossing occurring at only one point. The prediction based upon the postulated relationship between *D* and *H* is that the frequency curves will show an increasing divergence with training. The present data, perhaps because of the small number of cases at the higher MAS *A*-score levels, do not confirm this prediction. Confirmation on this particular point, based upon a substantial number of cases, has been reported (Spence, 1958).

SUMMARY

Fifty-eight *Ss* at various MAS *A*-score levels were submitted to an eyelid conditioning procedure in which no verbal ready signal was employed. Level of conditioning over 80 trials was found to be an increasing function of MAS *A* scores; the correlation between total number of CRs and *A* scores was linear and significant ($r = .40$).

The results are interpreted to be consistent with the assumptions that drive level is a function of emotionality of human *Ss*, as measured by the MAS *A* scale, and that the level of eyelid conditioning in turn is a function of drive level.

REFERENCES

BARON, M. R., BILLEY, J. J., & CONNOR, J. P. Verbal performance and eyelid conditioned responses with various levels of anxiety. *Amer. Psychologist*, 1958, **13**, 401–402. (Abstract)

GOODRICH, K. P., ROSS, L. E., & WAGNER, A. R. Performance in eyelid conditioning following interpolated presentations of the UCS. *J. exp. Psychol.*, 1957, **53**, 214–217.

HULL, C. L. *Principles of behavior*. New York: Appleton-Century, 1943.

RUNQUIST, W. N., & ROSS, L. E. The relation between physiological measures of emotionality and performance in eyelid conditioning. *J. exp. Psychol.*, 1959, **57**, 329–332.

RUNQUIST, W. N., & SPENCE, K. W. Performance in eyelid conditioning related to changes in muscular tension and physiological measures of emotionality. *J. exp. Psychol.*, 1959, **58**, 417–422.

SPENCE, K. W. Learning and performance in eyelid conditioning as a function of intensity of the UCS. *J. exp. Psychol.*, 1953, **45**, 57–63.

SPENCE, K. W. *Behavior theory and conditioning*. New Haven: Yale Univer. Press, 1956.

SPENCE, K. W. An emotionally based theory of drive (*D*) and its relation to performance in simple learning situations. *Amer. Psychologist*, 1958, **13**, 131–141.

SPENCE, K. W., & BEECROFT, R. S. Differential conditioning and level of anxiety. *J. exp. Psychol.*, 1954, **48**, 399–403.

SPENCE, K. W., & FARBER, I. E. Conditioning and extinction as a function of anxiety. *J. exp. Psychol.*, 1953, **45**, 116–119.

SPENCE, K. W., & TAYLOR, J. A. Anxiety and strength of the US as determiners of

the amount of eyelid conditioning. *J. exp. Psychol.*, 1951, **42**, 183–188.

SPENCE, K. W., & TAYLOR, J. A. The relation of conditioned response strength to anxiety in normal, neurotic, and psychotic subjects. *J. exp. Psychol.*, 1953, **45**, 265–272.

TAYLOR, J. A. The relationship of anxiety to the conditioned eyelid response. *J. exp. Psychol.*, 1951, **41**, 81–92.

TAYLOR, J. A. A personality scale of manifest anxiety. *J. abnorm. soc. Psychol.*, 1953, **48**, 285–290.

(Received October 16, 1959)

Journal of Experimental Psychology
1960, Vol. 60, No. 5, 314–317

LOCAL VARIATIONS IN THE MAGNITUDE OF A FIGURAL AFTEREFFECT

DONALD R. MEYER, SEISOH SUKEMUNE,[1] AND ROGER MYERS[2]

Ohio State University[3]

This experiment involves the measurement of magnitude of figural displacement as a function of the positions of the figures in the visual field. It was undertaken as a test of the assumption that the visual cortex is involved in the basic interactions that give rise to this phenomenon. Both the Köhler-Wallach (1944) and the Osgood-Heyer (1952) theory have this point in common, despite their markedly divergent postulates with respect to nature of the process. We will not be here concerned with the basic merits of the other features of these theories, but with a prediction that is reasonably derived from either one.

A specific locus clearly cannot be as readily established as discounted; and indeed, so long as the approach is indirect, the worth of an hypothesis derives from its ability to resist attempts at its discredit. The cortical hypothesis, as first proposed by Köhler (1940), has thus far been a viable concept; it was first arrived at with the demonstration that displacement equally obtains regardless of whether figures are examined with the same eye or with both eyes. This suggests involvement of central mechanisms and, if one considers classic pathways, the independence of the systems for the two eyes in the primates at the level of geniculates (cf. Walls, 1953) seemingly implies that the visual cortex is the first of possibilities. This presumes, of course, that the classic pathways are involved in such perceptions, and that we can discount crossed peripheral events of the kind that quite conceivably could be mediated by a central, extraclassical control of sensitivities of retinal networks by the visual efferents (Granit, 1955).

A further test of Köhler's hypothesis is based upon the supposition that displacement should be governed by the retinal, and not apparent, sizes of the figures. Here the works of Prentice (1950) and of Hochberg and Bitterman (1951) have failed to make the possible negation; the later work of Sutherland (1954), however, may be taken as keeping this an open question. The present work is similar to these approaches, and has been derived from the well-established inequalities of central visual representation for regions of the retina which lie more centrally and more peripherally.

The Osgood-Heyer theory leads us to expect displacement to reflect these differences, for their interpretation of displacement is in terms of processes that also figure prominently in the theory of acuity (Marshall & Talbot, 1942). From this point of view displacement and acuity should prove to be inversely related; sharp and separate excitation patterns cannot yield substantial successive interactions.

The general expectation from the Köhler-Wallach theory is a little less explicit, the problem being that inspections carried out in different portions of the visual field presumably induce satiations that are added to

[1] Now at the University of Hiroshima.
[2] Now at the University of North Dakota.
[3] From the Laboratory of Comparative and Physiological Psychology.

a gradient of permanent satiation. However, the mechanisms for establishment of the latter, baseline condition lead us to expect corresponding differences in temporary satiations. If such satiations are considered independent of the resting gradient, then the same prediction for relation of displacement to locus is obtained as before.

METHOD

The measures were obtained from two trained *Ss*, and materials were presented in a modified tachistoscope of the Harvard type. The figures and effects were similar to those presented and discussed by Köhler (1951, Fig. 3, p. 203). Each inspection (I) figure consisted of a black open circle on a large white background, and its size was such that it subtended 2° of visual angle (21 mm. at a distance of 611 mm.). Three such I figures were employed; these differed only in that the circles were so placed that their centers and a central fixation point subtended either 2°, 4°, or 6° of visual angle. The center points of circles and the fixation point were horizontally disposed, and the circles invariably appeared on the left side of the card.

Each test (T) figure consisted of a pair of circles, the centers of which were equidistant from a central fixation point identical to that employed for the I figures. These T figures were displayed on three series of five test cards each. Each pair included one "affected" circle or circle to be subject to displacement; this was arranged to be concentric with the circle on one inspection card, and was of such a size that it subtended 1° of visual angle. Corresponding "unaffected" circles in a series for a given lateral position were graduated downward from equality in steps of 1 mm. Each step thus corresponded fairly closely to a 10% reduction. The fixation points were 2 mm. in diameter; lines of all I and T circles were 2 mm. wide.

It will be apparent that the general approach was that of measurement of shrinkage in affected test circles by the constant stimulus method. The choice of the 2:1 size ratio for I and T figures was determined by results of Japanese experiments that have been reviewed in English by Sagara and Oyama (1957). Background luminance for all I figures was 6.4 ft-L, that for T figures was 7.1 ft-L.

The use of this approach has one important complication, namely that the judgments may be subject to bias by the half-meridional differences described by Brown (1953). Ancillary judgments were made with cards designed to assess this situation; these differed from the experimental test cards in that the series of right-hand circles were pivoted on physical equivalence with the left-hand circles. The method was as in the study proper, but of course without inspections; the results obtained led us to dismiss resting differences as trivial.

For each experimental presentation, an I figure was exposed for 20 sec. It was viewed binocularly by *S*, who maintained a central fixation. At the end of the inspection time, a pair of T figures of equal eccentricity replaced the I figure for a .5-sec. period. The *S* was required to make a judgment as to which of the two T circles was the larger. The 15 possible conditions were presented in random order, and each condition was judged 100 times by each of the two *Ss*.

RESULTS

The main conditions and results are summarized in Fig. 1. For both *Ss*, the extent of shrinkage of the affected T circle—as determined by the diminution which would be required for matching by an unaffected circle—proved to be a function of the eccentricities of the figures being examined. Overall differences, by Friedman's analysis of variance for ranks, yielded *P* values beyond .01 and .03 for individual *Ss*. For Subject SS the obtained effect ranged from 10% to slightly over 20%, and the increments appeared in regular relation to retinal position. Subject RM's shrinkages were larger, ranging upward to a third, but with an apparent though unreliable reversal of the outer functions.

The question arose as to whether these results could be related in a simple way to individual differences in variations of acuity with retinal position. The data required for this assessment were obtained with a zoom system acuity meter utilizing parallel bars. For the present pur-

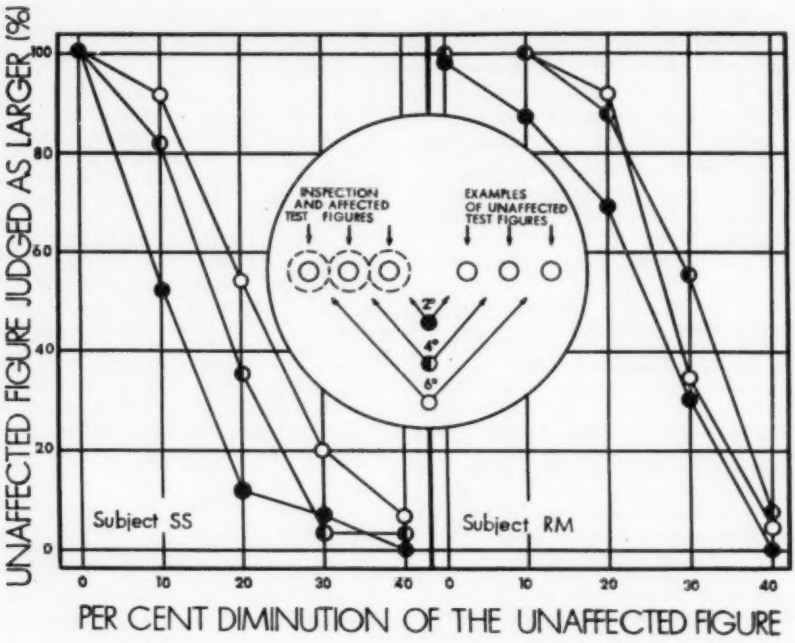

FIG. 1. Summary of judgments in relation to experimental conditions.

pose, these are best expressed in completely relative terms; accordingly, for each *S* the reciprocal of distance required for resolution of the bars at 2° was taken as a score of 100. For SS relative acuity scores for bars at 4° and 6° were 86 and 72, respectively; for RM these same scores were 85 and 71. We were thus required to dismiss the concept under test, and can offer no alternative.

In general, we conclude from this experiment that magnitude of figural displacement varies systematically with retinal position and, by implication with locus of related central processes. Displacement thus reflects a well-known property of visual neural organization, being greatest when the figures fall where acuity is poorest. For one *S* this relation is precise; for the other we can merely say that there exists a difference between the more central and the more peripheral conditions. This result, while not itself sufficient, is consistent with the supposition that the visual cortex is the site of interactions underlying such phenomena.

SUMMARY

Measurements were made with two *Ss*, by a constant stimulus method, of magnitude of figural displacement as a function of position of the figures. Generally, displacement is greatest when the interacting figures are peripherally placed within the central visual field. This result is consistent with the view, common to theories of displacement, that the visual cortex is the locus of the processes involved in this effect.

REFERENCES

BROWN, K. T. Methodology for studying figural after-effects and practice effects in the Mueller-Lyer illusion. *Amer. J. Psychol.*, 1953, **66**, 629–634.

GRANIT, R. Centrifugal and antidromic effects on ganglion cells of the retina. *J. Neurophysiol.*, 1955, **18**, 388–411.

HOCHBERG, J. E., & BITTERMAN, M. E. Figural after-effects as a function of the retinal size of the inspection figure. *Amer. J. Psychol.*, 1951, **64**, 99–102.

KÖHLER, W. *Dynamics in psychology.* New York: Liveright, 1940. Pp. 86–87.

KÖHLER, W. Relational determination in perception. In L. A. Jeffress (Ed.), *Cerebral mechanisms in behavior.* New York: Wiley, 1951. Pp. 200–230.

KÖHLER, W., & WALLACH, H. Figural after-effects: An investigation of visual processes. *Proc. Amer. Phil. Soc.*, 1944, **88**, 269–357.

MARSHALL, W. H., & TALBOT, S. A. Recent evidence for neural mechanisms in vision leading to a general theory of sensory acuity. In H. Klüver (Ed.), *Visual mechanisms.* Lancaster, Pa.: Cattell, 1942. Pp. 117–164.

OSGOOD, C. E., & HEYER, A. W. A new interpretation of figural after-effects. *Psychol. Rev.*, 1952, **59**, 98–118.

PRENTICE, W. C. H. On the size of figural after-effects with varying distances. *Amer. J. Psychol.*, 1950, **63**, 589–593.

SAGARA, M., & OYAMA, T. Experimental studies on figural aftereffects in Japan. *Psychol., Bull.*, 1957, **54**, 327–338.

SUTHERLAND, N. S. Figural after-effects, retinal size, and apparent size. *Quart. J. exp. Psychol.*, 1954, **6**, 35–44.

WALLS, G. L. *The lateral geniculate nucleus and visual histophysiology.* Berkeley and Los Angeles: Univer. California Press, 1953. Pp. 12–13.

(Received October 16, 1959)

Journal of Experimental Psychology
1960, Vol. 60, No. 5, 318–322

TWO-DIMENSIONAL TRACKING WITH IDENTICAL AND DIFFERENT CONTROL DYNAMICS IN EACH COORDINATE

RUBE CHERNIKOFF, JOHN W. DUEY, AND FRANKLIN V. TAYLOR[1]

United States Naval Research Laboratory

In many practical man-machine control devices the human is required to perform different tracking tasks in various coordinates. In an airplane, for example, the pilot must supply one transfer function while controlling the heading of his craft through left and right motions of his stick, and a different transfer function while controlling pitch with fore and aft stick motions. The necessity for these different modes of action arises from differences in the dynamic properties of the two control loops. These dissimilar dynamics, in turn, result from the physics of regulated flight.

The purpose of the present investigation was to determine the effect upon tracking performance produced by using various combinations of dynamics in the X and Y coordinates of a two-dimensional tracking system. We were interested in discovering what the manipulation of the dynamics in one coordinate would do to tracking performance in the other.

Three sets of dynamics were employed: (*a*) Position Control (zero-order system). With this, force applied to the joystick is translated directly into a displacement of the displayed marker. (*b*) Rate Control (first-order system). An integrator inserted beyond the control causes the marker to move at a velocity which is proportional to the force applied to the control. (*c*) Acceleration Control (second-order system). Two integrators are inserted beyond the control. As a result, any application of force to the control produces a proportional acceleration of the marker.

These three types of dynamics each require a characteristically different control response from *S*. With position control, he simply displaces his control an appropriate amount in the appropriate direction. With rate control, *S* displaces the control until the generated velocity moves the marker to the intended position. He then makes a second movement to return the control to center, removing the velocity. With acceleration control, *S* displaces the control until the generated acceleration of the dot brings it near the desired position. He then must move the control in the reverse direction by an amount appropriate to remove the generated acceleration. He finally makes a third movement to return the control to center. Thus, adding integrators may result in increasing the complexity of the required control movements.

In this study, the three dynamics were paired in all combinations on the X and Y control coordinates.

METHOD

Apparatus.—A block diagram of the two-coordinate tracking system is shown in Fig. 1. The two switches, X_A and Y_A, are used to select the type of control dynamics to be used in each coordinate. When either switch is set on Contact 1, position control tracking is obtained. Setting a switch at Contact 2 provides rate control, while for acceleration control the switch is set at Contact 3. Each switch can be set separately so that all combinations of dynamics in X and Y are available. Switches X_B and Y_B are used to eliminate tracking in one coordinate. This

[1] The authors are indebted to H. P. Birmingham for his provocative suggestions relative to this problem and for assistance in apparatus design.

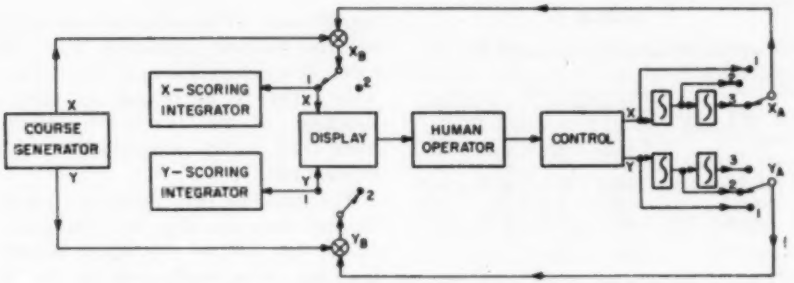

FIG. 1. Block diagram of the two-coordinate tracking system.

is done by setting a switch at Contact 2. Placing either switch at Contact 1 permits control in that coordinate, and provides a course input to the display.

The course input for each of the coordinates consisted of a single sine wave, differing in frequency and amplitude. In the X coordinate the frequency was 1.9 cycles/min., while in the Y coordinate a frequency of 3.15 cycles/min. was used. The amplitudes of the sine waves were inversely proportional to their frequencies. The course voltages to the display were sufficient to generate amplitudes of ±5 in. for the X coordinate and ±3 in. for the Y coordinate.

The display was a 5-in., short-persistence cathode ray tube (CRT). A compensatory type of tracking display was used. That is, variations in the course moved a marker dot away from a 1/8-in. center reference circle on the CRT, and *S* made the appropriate control movements to return the dot to center. The control consisted of a spring-restrained joystick which was free to move at any angle so that the marker dot could be moved to center directly from any position on the display. The joystick-to-marker sensitivity for both coordinates was such that a 10° displacement of the stick produced a marker change on the CRT face of 1 in. with position control, 1 in./sec. with rate control, and 1 in./sec.^{2} with acceleration control.

The control dynamics were obtained by means of electronic analog computer elements. The gains of the integrators used for the rate control and the acceleration control systems were set at unity for each course frequency employed. This provided equivalent control system gain for each of the dynamics used.

Tracking error, consisting of the difference between the course input and system output, was scored separately for each coordinate by use of two identical electronic scoring integrators. The error produced in each coordinate was rectified and integrated over the scoring duration by a scoring system composed of electronic computer elements operating from the error signal to the display. Integrated error scores were read directly from a voltmeter after each trial and represent the average error for the trial. Since the time of each tracking trial is held constant, integrated error is proportional to average error.

Procedure.—Tracking error scores were first obtained for the three types of dynamics tracked in a single coordinate, both in X and in Y. Following this, the same *S*s were given all nine combinations of dynamics in two-coordinate tracking. The six conditions of one-coordinate tracking were presented for nine sessions. A session consisted of three 1-min. trials on each condition, with the error accumulated during the last 45 sec. of each trial scored by the error integrator. With two-coordinate tracking, all nine pairings of dynamics in the X and Y coordinates were run for 20 sessions, three trials per session. Six naval enlisted men served as *S*s. The order in which the experimental conditions were run was randomized for each *S* during each session.

RESULTS AND DISCUSSION

Table 1 presents the results for the single-coordinate tracking. The entries are the means and *SD*s of the error scores for the six *S*s recorded during the last three sessions of tracking. Comparisons between coordinates, in which different course inputs were used, showed that the low-frequency, high-amplitude course in the X coordinate was significantly

TABLE 1

SINGLE COORDINATE TRACKING DATA

Dynamics	X Coordinate		Y Coordinate	
	Mean	*SD*	Mean	*SD*
Position	80.2	15.9	97.3	11.1
Rate	24.2	8.9	34.6	8.3
Acceleration	44.4	10.1	49.2	16.7

better ($P = .05$) than the high-frequency, low-amplitude course in Y for position and rate control dynamics. No significant difference in performance was found for acceleration control between coordinates. However, all differences in tracking error among the three types of control dynamics were significant at the 5% level. In this investigation the Wilcoxon (1949) paired-replicates test was used in all tests of statistical significance. The order of proficiency for the various dynamics was: (*a*) rate, (*b*) acceleration, and (*c*) position. This order is, of course, in part a function of the particular equipment and inputs employed in this experiment.

The data for two-coordinate tracking are shown in Fig. 2. The graph on the left half of this figure depicts tracking error measured in the X coordinate. On the abscissa are shown the dynamics in Y which were paired with the various dynamics in X. The tracking error in X is used in plotting the parameters shown. Similarly, the graph on the right half of the figure shows the parameters obtained by plotting the tracking error measured in the Y coordinate. It should be noted that the components of error along the X and Y coordinates were scored simultane-

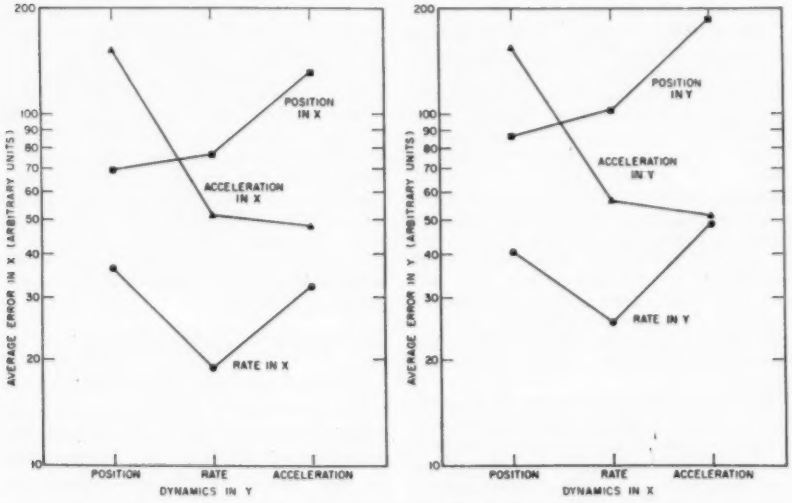

FIG. 2. Two-coordinate tracking data. The graph on the left depicts tracking error measured in the X coordinate, when paired with the various dynamics in Y shown on the abscissa. Similarly, the graph on the right shows the parameters obtained by plotting tracking error measured in the Y coordinate.

TABLE 2

P VALUES FOR STATISTICAL COMPARISONS BETWEEN DYNAMICS FOR TWO COORDINATE TRACKING

Comparisons	Coordinate Scored	
	X	Y
Pp vs. Pr	.05	.05
Pp vs. Pa	.05	.05
Pr vs. Pa	.05	.05
Rp vs. Rr	.05	*ns*
Rp vs. Ra	*ns*	*ns*
Rr vs. Ra	.05	.05
Ap vs. Ar	.05	.05
Ap vs. Aa	.05	.05
Ar vs. Aa	*ns*	*ns*

Note.—The letters P, R, and A in the Comparisons column represent the position, rate, and acceleration dynamics, respectively. The capital letters indicate the dynamics which are compared, while the small letters refer to the dynamics used in the other coordinate. The columns X and Y indicate the significance levels for each of the coordinates scored.

ously for the various pairings of dynamics, and the separation into the two graphs by coordinates is simply for convenience in presenting the data.

Table 2 presents the results of the various comparisons and indicates which are significant. For six *S*s, the Wilcoxon paired-replicates test provides for a confidence level no greater than 5%.

From Fig. 2 it can be seen that in all cases the minimum amount of error produced with each type of dynamics occurred when the dynamics in the two coordinates were the same. Increasing the difference in the number of integrators comprising the dynamics of each coordinate produces an increase in tracking error. However, from the curves of Fig. 2 it can be seen that a difference of one integrator between dynamics results in only a nominal increase in error, whereas a difference of two integrators between dynamics contributes a considerable increase in tracking error.

To provide a different tracking task in each dimension, the frequencies and amplitudes used as inputs to each coordinate differed in the present experiment. In a preliminary study, the same course was used for each coordinate. The results from the preliminary study were very similar to the findings for the present investigation. Thus, the use of a different tracking task on each coordinate provides some generality to the results and, at the same time, does not affect the basic relationships found by manipulating control dynamics.

The results of this study demonstrate clearly that tracking performance in a given coordinate is influenced by the dynamics used in the second coordinate. The findings are consonant with the hypothesis that the human has a limited information handling capacity or "bandwidth" (Birmingham & Taylor, 1954) and that the greater the difference between control tasks in the two coordinates the less the capacity which remains available for the performance of either. Apparently, the act of shifting from one transfer function to another itself consumes bandwidth which must be subtracted from that available for the performance of the two tracking tasks. As more and more of the human's capacity is, thus, dissipated with increasing dissimilarity between tasks less and less remains to be devoted to tracking, with the result that performance accuracy drops in both coordinates.

SUMMARY

In many complex tracking devices the control dynamics may differ in the various coordinates, thus requiring *S* to supply a different transfer function in each tracking dimension. The purpose of the present investigation was to determine the effect upon tracking produced by using various pairings of position, rate, and acceleration control dynamics in the X and Y coordinates of a two-dimensional tracking system.

The results indicate that in all cases minimum error occurred when the dynamics

in the two coordinates were the same and that error increased as the dissimilarity between the dynamics became greater.

These findings are interpreted as suggesting that the human has a limited information handling capacity, that some of this capacity is used to change from one task to the other, and that the greater the difference between control tasks the less the capacity which remains for the performance of either.

References

BIRMINGHAM, H. P., & TAYLOR, F. V. A design philosophy for man-machine control systems. *Proc. Inst. Radio Engrs., NY*, 1954, **42**, 1748–1753.

WILCOXON, F. *Some rapid approximate statistical procedures.* Stamford, Conn.: American Cyanamid, 1949.

(Received October 26, 1959)

Journal of Experimental Psychology
1960, Vol. 60, No. 5, 323–326

INSTRUMENTAL ESCAPE PERFORMANCE AS A FUNCTION OF THE INTENSITY OF NOXIOUS STIMULATION

MILTON A. TRAPOLD

State University of Iowa

AND

HARRY FOWLER [1]

University of Pittsburgh

While a large amount of work has been done on the role of the primary motivational variables (e.g., food or water deprivation) in the area of instrumental appetitive conditioning, few systematic data are available on the role of motivational variables such as electric shock in instrumental escape conditioning. Moreover, as Spence (1956) has pointed out, those data which are available are quite inconsistent as to the form of the functional relationship between asymptotic escape performance and shock intensity. As these data are extensively discussed by Spence (1956) only brief mention of the inconsistencies will be made here.

While Amsel (1950) found that a group of rats receiving a strong shock ran faster than a group receiving a weak one, Campbell and Kraeling (1953) report no differences in asymptotic performance level for *Ss* running down a straight alley charged with shock intensities of either 200, 300, or 400 v. to an uncharged goalbox. Using much lower shock levels, Ketchel, in an unpublished study cited by Spence (1956), found that *Ss* escaping a 36-v. shock performed at a higher level than those escaping an 11.6-v. shock. However, Ketchel's weak-shock group showed little or no increase in performance over trials, raising the possibility that this weak shock was subthreshold and as such provided no motivation.

In view of the inconsistencies among these various results, the present study seeks to provide more extensive data on the effect of shock intensity in instrumental escape learning. A wide range of shock intensities is examined so as to provide a more general picture of the function relating escape performance to shock intensity.

METHOD

Subjects.—The *Ss* for this experiment were 30 male albino rats, all of which had served earlier in a food-reinforced straight alley experiment entailing the administration of a few brief and relatively weak shocks. Fifteen of these *Ss* were approximately 300 days old at the time of the present experiment, while the other half were approximately 180 days old.

Apparatus.—A straight alley 68 in. long, 3 in. wide, and 4 in. high served as the experimental apparatus. The alley was constructed of two L-shaped strips of galvanized sheet metal, and was covered with a transparent Plexiglas top. One of the sheet metal strips formed one side and half the floor of the apparatus; the other strip formed the other side and other half of the floor. A 1-in. gap separated the two strips at the floor.

The alley was separated into three sections; a startbox 8 in. long, an alley section 48 in. long, and a goalbox 12 in. long. These sections were separated by upward acting guillotine doors. Raising the start door operated a microswitch which started a Standard Electric clock and closed a shock circuit to the start box and alley. Breaking a photobeam 6 in. inside the alley stopped the first clock and started a second which in turn was stopped when *S* broke a photobeam

[1] This experiment was conducted at Yale University while H. Fowler held a predoctoral fellowship from the National Institute of Mental Health, Bethesda, Maryland.

2 in. inside the goalbox. Breaking this second photobeam also interrupted the shock circuit to the startbox and alley. The two time measures were recorded to the nearest .01 sec. and converted to speed scores (100/time). Hereafter these measures will be referred to as starting and running speed, respectively.

A single matched impedance 60 cps ac shock source provided shock for both the startbox and alley. In this shock system, the galvanized metal strips forming the halves of the alley were connected across the output of a power transformer through a voltage divider which provided variation in the applied voltage. In addition, a .25 megohm resistor was placed in series with *S*. The *S* received shock as long as its feet touched both of the metal strips forming the halves of the alley. The narrowness of the alley prevented *S* from running along only one side of the alley.

Procedure.—The *S* was placed in the startbox with the startbox door closed and the goalbox door open. When *S* oriented toward the goal, the startbox door was opened and *S* could run to the goal end of the alley. When *S* entered the goalbox, the goalbox door was closed and *S* was detained in the goal box for 20 sec. During this time, *E* recorded the data from that trial and reset the apparatus. At the end of the 20 sec., *S* was removed from the goalbox and immediately returned to the startbox for the beginning of the next trial. In this manner each *S* received a total of 20 highly massed trials.

Experimental design.—The 30 *Ss* were randomly assigned to five experimental groups in such a way as to randomize past experimental history as well as all other extraneous factors, with the restriction that each group contain 3 *Ss* from each of the two age populations. This restriction permitted isolation of the variance due to the age variable in the statistical analyses.

Each group of *Ss* received a different intensity of shock in the startbox and alley; shock was absent for all groups in the goalbox. The shock intensities used were 120, 160, 240, 320, and 400 v., as measured directly across the output of the voltage divider. It was originally planned to include a group at 80 v., but preliminary investigation showed that *Ss* would not run at this low voltage.

Results

Mean speed curves for each group are presented in Fig. 1 and 2. Figure 1 shows the learning curves over the entire course of the experiment for both the starting and running speed measures. Figure 2 shows for these two measures the relationship between mean speed on the last 8 trials and shock intensity.

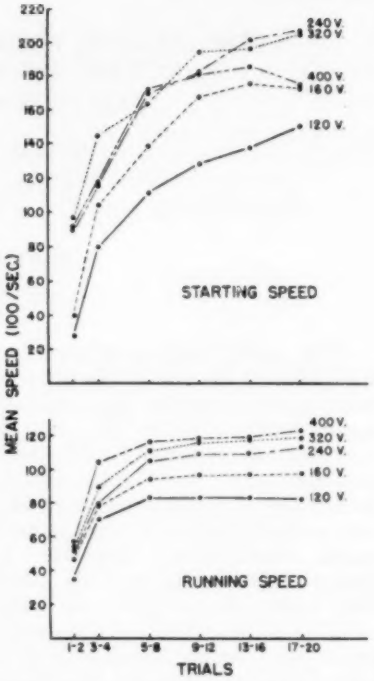

FIG. 1. Learning curves for the various shock intensity groups for both the starting speed and running speed measures.

The results of an analysis of variance of mean running speed over the last 8 trials showed that the young *Ss* ran slightly, but not significantly, faster than the old *Ss* ($F = 2.97$; $df = 1, 20$; $P = .10$). The F ratio for the interaction of shock intensity and age was less than 1. Because of this negligible interaction, and a similar result in the case of starting speed, the age variable, which is

inconsequential to the main purposes of this experiment, was ignored in Fig. 1 and 2.

The overall F for the different shock intensity conditions approached but also failed to meet conventional levels of statistical significance ($F = 2.26$; $df = 4, 20$; $P < .10$). However, inspection of the running speed curve in Fig. 2 shows that the curve is negatively accelerated in form, with the three highest shock intensity groups being very close to one another in mean speed. The approximate equality in performance of these three groups would tend to reduce the overall shock intensity effect, thus producing the nonsignificant overall F. A trend analysis of these data, by orthogonal polynomial comparisons, indicates that running speed was significantly related to log shock intensity ($F = 8.53$; $df = 1, 20$; $P < .01$), with the F for the residual variance after extraction of this log linear component being less than 1.

Figure 2 also shows that the form of the relationship between starting speed and shock intensity is quite different from that in the case of running speed. Starting speeds first increase with increasing shock intensity, but beyond 240 v., further increases in shock intensity lead to progressive decrements in this measure. The results of an analysis of variance of mean starting speeds on the last 8 trials showed that the overall effect of shock intensity was significant ($F = 1.61$; $df = 4, 20$; $P < .005$). However, the main effect of age was again not significant, and the Shock × Age interaction again yielded an F less than 1. Analysis of the trends within these data by orthogonal polynomials showed that the functional relationship between starting speed and shock intensity contained both a significant linear component ($F = 9.29$; $df = 1, 20$; $P < .01$) and a significant quadratic component ($F = 13.18$; $df = 1, 20$; $P < .005$). The residual variance after extraction of these two components was negligible ($F < 1$).

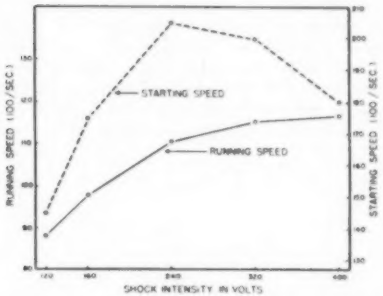

FIG. 2. Mean performance over the last 8 trials as a function of shock intensity.

DISCUSSION

In total the results show that, over the range of shock intensities used in the present study, running speed is an increasing, negatively accelerated function of shock intensity, while starting speed is nonmonotonically related to shock intensity, first increasing and then decreasing with increasing shock intensities.

The results of the running speed measure are in line with the results of Amsel (1950) and Ketchel (Spence, 1956), but not those of Campbell and Kraeling (1953). While it is true that part of the inconsistency in these previous studies may have arisen from the use of different procedures and types of shock, the fact that the present study used a shock source, alley, and general procedure as close as possible to that used by Campbell and Kraeling makes this interpretation unlikely. Perhaps a more reasonable interpretation is that Campbell and Kraeling found no differences in asymptotic performance because all of their shock intensity groups were far out on the relatively flat portion of the curve relating speed

to shock intensity. If it is also true that starting speed is, with this general procedure and apparatus, nonmonotonically related to shock intensity, then this factor could also have contributed to their lack of differences, as the speed measure in the Campbell and Kraeling experiment included both the starting and running speed measures of the present experiment.

Summary

Five groups of rats were trained to escape shocks of 120, 160, 240, 320, and 400 v. in a straight alley by running to an uncharged goalbox. Running speed was found to be an increasing, negatively accelerated function of shock intensity, but starting speed was nonmonotonically related to shock intensity, first increasing and then decreasing with increasing shock intensity. These results are discussed in relation to inconsistencies among the findings of several other studies pertaining to the empirical relationship between shock intensity and performance in escape conditioning.

References

Amsel, A. The combination of a primary appetitional need with primary and secondary emotionally derived needs. *J. exp. Psychol.*, 1950, **40**, 1–14.

Campbell, B. A., & Kraeling, D. Response strength as a function of drive level and amount of drive reduction. *J. exp. Psychol.*, 1953, **45**, 97–101.

Spence, K. W. *Behavior theory and conditioning.* New Haven: Yale Univer. Press, 1956.

(Received October 30, 1959)

Journal of Experimental Psychology
1960, Vol. 60, No. 5, 327

SUPPLEMENTARY REPORT: DELAY OF KNOWLEDGE OF RESULTS, KNOWLEDGE OF TASK, AND INTERTRIAL INTERVAL

M. RAY DENNY, MARVEL ALLARD, EUGENE HALL, AND MILTON ROKEACH

Michigan State University

Bilodeau and Bilodeau (1958) present a convincing array of data in delay of knowledge of results (KR) experiments that the intertrial interval rather than KR delay is the critical variable (better learning with shorter interval). In one study that they cite (Saltzman, Kanfer, & Greenspoon, 1955) the intertrial interval varied directly with KR delay and on either count one might have expected differences in learning, but the results were negative. One procedural difference between the studies yielding positive intertrial interval results and the study by Saltzman et al. (1955) is that in the latter study *S* knew, in a clear mediational sense, what the task was, i.e., to draw a 3-in. line blindfolded. The present hypothesis is that if *S* were unclear or even misled about the task (response to be learned is unidentified) that the intertrial interval and perhaps KR delay would become relevant. Thus the present study is an exact replication of the Saltzman et al. (1955) study with the exception that *S* was instructed to draw a line 60 "glubs" long (actually 3 in. long).

Method.—All *Ss* ($N = 96$) were from introductory psychology classes. As in the Saltzman et al. study there were three experimental groups ($N = 24$ in each): 0-sec. KR delay; 10-sec. KR delay; 20-sec. KR delay. All three groups had a post-KR delay of 10 sec. prior to starting the next trial. Unlike the Saltzman et al. study, there was also a control group ($N = 24$), 0-sec. KR delay with a 30-sec. post-KR delay; thus two groups had identical intertrial intervals and differed maximally on KR delay, and two groups had identical KR delay and differed maximally on intertrial interval. The criterion of learning was drawing five consecutive lines between $2\frac{3}{4}$ in., and $3\frac{1}{4}$ in., or a maximum of 60 trials. All information given to *S* about the line drawn was with respect to "glubs" rather than inches.

Results and discussion.—The 0-sec., 10-sec., and 20-sec. delay groups learned in a median of 18.5, 25.5, and 41.0 trials, respectively, as compared with 28.0, 25.0, and 30.5 trials in the study by Saltzman et al. With the Mann-Whitney test, the 0-sec. and the 10-sec. delay groups were significantly better than the 20-sec. group ($z = 2.47$ and 2.39, $P = .02$), but the difference between the 0-sec. and 10-sec. groups was nonsignificant. The control group had a median of 34.5 trials and except for a few middle scores, had a distribution almost identical with the 20-sec. delay group, including 9 *Ss* who took 60 trials as compared with 7 in the 20-sec. delay group. The control was significantly worse than the 0-sec. delay group ($z = 2.53$, $P = .02$) but did not differ significantly from the other two.

Given the results of the control group the interpretation by Bilodeau and Bilodeau (1958) receives clear support. The intertrial interval confounded, as it may be, with post-KR delay appears critical *rather than* KR delay. In addition, there was clear support for the hypothesis that previous experience in conjunction with knowledge of the response to be learned will diminish the importance of intertrial interval as a variable. Presumably if control groups had been incorporated in the study by Saltzman et al. (1955), no additional information would have been gained; all groups would have performed alike because the positive transfer effect from previously using and judging inches probably would have bridged the intertrial interval.

REFERENCES

BILODEAU, E. A., & BILODEAU, I. McD. Variation of temporal intervals among critical events in five studies of knowledge of results. *J. exp. Psychol.*, 1958, **55**, 603–612.

SALTZMAN, I. J., KANFER, F. H., & GREENSPOON, J. Delay of reward and human motor learning. *Psychol. Rep.*, 1955, **5**, 139–152.

(Received April 9, 1960)

ERRATUM

The figure below was to have appeared as Part B of Fig. 1 in the article "Successive Acquisitions and Extinctions as Related to Percentage of Reinforcement" by Glen D. Jensen and John W. Cotton (*J. exp. Psychol.*,1960, 60, 41–49). The figure is necessary in connection with the text on pp. 42–43.

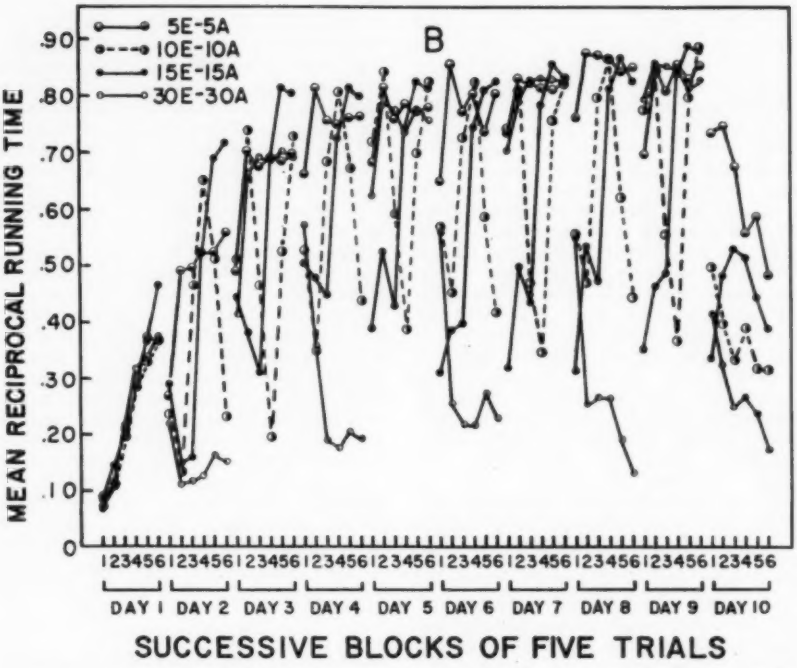

FIG. 1. (Part B) Mean reciprocal running time in seconds as a function of trials. Trials are grouped in blocks of five. Day 10 is an extinction session for all groups. The coded parameter is the number of extinction trials (E) and the number of acquisition trials (A) in alternating sets of extinction and acquisition trials on Days 2 through 9.